

Genre analysis
of online encyclopedias
The case of Wikipedia

Anna Tereszkievicz

Genre analysis of online encyclopedias

The case of Wikipedia



Wydawnictwo Uniwersytetu Jagiellońskiego

Publikacja dofinansowana przez Wydział Filologiczny Uniwersytetu Jagiellońskiego ze środków wydziałowej rezerwy badań własnych oraz Instytutu Filologii Angielskiej

PROJEKT OKŁADKI

Bartłomiej Drosdziok

Zdjęcie na okładce: *Łukasz Stawarski*

© Copyright by Anna Tereszkievicz & Wydawnictwo Uniwersytetu Jagiellońskiego
Wydanie I, Kraków 2010
All rights reserved

Książka, ani żaden jej fragment nie może być przedrukowywana bez pisemnej zgody Wydawcy. W sprawie zezwoleń na przedruk należy zwracać się do Wydawnictwa Uniwersytetu Jagiellońskiego.

ISBN 978-83-233-2813-1



www.wuj.pl

Wydawnictwo Uniwersytetu Jagiellońskiego
Redakcja: ul. Michałowskiego 9/2, 31-126 Kraków
tel. 12-631-18-81, 12-631-18-82, fax 12-631-18-83
Dystrybucja:
tel. 12-631-01-97, tel./fax 12-631-01-98
tel. kom. 0506-006-674, e-mail: sprzedaz@wuj.pl
Konto: PEKAO SA, nr 80 1240 4722 1111 0000 4856 3325

Table of Contents

Acknowledgements	9
Introduction	11
Materials and Methods.....	14
1. Genology as a study	15
1.1. Schools of genre analysis.....	15
1.2. Genre – characteristics of the concept	16
1.3. Genre evolution.....	17
1.4. Discourse and genre competence.....	18
1.4.1 Discourse competence.....	18
1.4.2 Generic competence	19
2. Genres on the web	21
2.1. Genre theory on the web	21
2.2. The concept of web genre	22
2.3. Genres on the web – taxonomy	23
2.4. Genre hybridization and individualization.....	26
2.5. Web genre evolution	27
3. Encyclopedia as a genre.....	29
3.1. Encyclopedia as a genre of information.....	29
3.2. Genre characteristics of an encyclopedia.....	30
3.2.1. Function.....	32
3.2.2. Content	33
3.2.2.1. Authorship	33
3.2.2.2. Scope	33
3.2.2.3. Adjuncts.....	34
3.2.3. Form	35
3.2.3.1. Density.....	35
3.2.3.2. Article structure	35
3.2.3.3. Classification	35
3.2.3.4. Cross-references	37
3.2.3.5. Discourse properties	38
4. Encyclopedias on the web.....	41
4.1. Online reference works.....	41

4.2. Typology of online encyclopedias	42
4.2.1. Reproduced encyclopedias.....	43
4.2.2. Adapted encyclopedias.....	45
4.2.2.1. Lower adapted encyclopedias.....	45
4.2.2.2. Advanced encyclopedias.....	47
4.3. Emergent encyclopedias	51
4.3.1. User-created portals.....	51
4.3.2. Resource pages.....	57
4.4. Emerging encyclopedias	59
5. Wikipedia as a genre.....	65
5.1. Wikipedia – introduction to the project	66
5.1.1. Wikipedia mechanism	66
5.1.2. Dimensions of the project	67
5.1.3. Legal issues	69
5.2. Wikipedia versus traditional encyclopedia	69
5.2.1. Function of Wikipedia.....	70
5.2.2. Content	70
5.2.2.1. Authorship	70
5.2.2.1.1. Wikipedia community.....	72
5.2.2.2. Scope	73
5.2.2.3. Adjuncts.....	74
5.2.2.4. Genre syncretism	81
5.2.3. Form	82
5.2.3.1. Density	82
5.2.3.2. Article structure.....	82
5.2.3.3. Classification.....	87
5.2.3.4. Cross-references.....	90
5.2.4. Functionality.....	92
5.2.5. Meeting the quality standards.....	94
6. Wikipedia discourse features	99
6.1. The continuum of Wikientries.....	99
6.1.1. Conventional articles.....	100
6.1.2. Unconventional articles.....	103
6.1.2.1. Textual features.....	103
6.1.2.1.1. Textual organization – integration versus fragmentation.....	103
6.1.2.1.2. Heterogeneity.....	108
6.1.2.1.3. Precision and exemplification.....	123
6.1.2.1.4. Repetitiveness	126
6.1.2.1.5. Informality	127
6.1.2.1.6. Personalization.....	130
6.1.2.2. Syntactic features.....	133
6.1.2.2.1. Lack of nominalization	133
6.1.2.2.2. Activation	134
6.1.2.2.3. Structure of sentences	135
6.1.2.2.4. Epanaphora	136
6.1.2.2.5. Informality	136
6.1.2.3. Lexical features.....	137
6.1.2.3.1. Authorial presence – subjectivity and interactivity.....	137
6.1.2.3.2. Informality	150
6.1.2.3.3. Terminology.....	152
6.1.2.3.4. Inverted commas.....	155

6.1.3. Anti-encyclopedia articles	158
6.1.3.1. Unintelligibility	158
6.1.3.2. Imprecision	160
6.1.3.3. Incoherence.....	161
6.1.3.4. Language mistakes	163
6.2. Wikidiscourse transformation areas	165
Conclusions	169
Bibliography	173
Sources.....	173
Online encyclopedias	173
References.....	175

Acknowledgements

I would like here to acknowledge the greatest help provided by Professor Elżbieta Mańczak-Wohlfeld. I am truly grateful for her guidance, valuable commentaries, patience and constant encouragement. I am truly indebted to Professor Elżbieta Chrzanowska-Kluczevska, Professor Ewa Willim and to Professor Andrzej Łyda for their insightful comments, helpful advice and their great support.

I am deeply beholden to the Jagiellonian University for honouring me with the Florentyna Kogutowska Fund, which enabled me to greatly advance my research and gave me the splendid opportunity to travel in the search of essential sources.

I am also grateful to Professor David Singleton and the Dublin Trinity College for accepting me as a Visiting Academic and offering all the great facilities of the College and the magnificent library collection at my disposal.

I also owe a lot to the community of Wikipedians for their great help in grasping the inner workings of the free encyclopedia.

Introduction

The World Encyclopedia would be the mental background of every intelligent man in the world. It would be alive and growing and changing continually, under revision, extensions and replacement from the original thinkers in the world everywhere. Every university and research institution should be feeding it. Every fresh mind should be brought into contact with its standing editorial organization. And on the other hand, its contents would be the standard source of material for the instructional side of school and college work, for the verification of facts and the testing of statements, everywhere in the world.(...) It would hold the world together mentally.

11

Wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them, ready to be dropped into the memex and there amplified.

The quotations above come from “World Encyclopedia”, by H.G. Wells (1938) and V. Bush’s “As We May Think” (1991), respectively. In his work, Wells pointed out that the dynamism of the changes in the 20th century creates a need for innovative types of encyclopedias. The Wellsian project, the *World Encyclopedia*, was planned as a synthesis of knowledge in the form of a collection of extracts, essays and reference lists. Bush, in his essay, proposes the implementation of a memex, a new device for communication and information storage, which was to facilitate the transfer and retrieval of knowledge, and in that way contribute to the compilation of modern encyclopedic sources. Both scholars may be considered as early visionaries of the potential of contemporary technology. It is only today, more than fifty years since the publications of these works, that we can observe the near accomplishment of their dreams.

Still, Wells and Bush were not the first to express a dream for a resource encompassing all knowledge. As Bolter (2001: 81) observes, the dream to compose a great book presenting the wisdom of mankind dates back as far as ancient Greece and Rome, continuing through mediaeval times and the Age of Enlightenment up to the present day. According to the scholar, this desire has evinced itself in the birth of the encyclopedia, a book designed to condense all knowledge.

The earliest encyclopedic texts, written mainly in the form of thematic treatises, constituted a means of describing the nature of the external world. During the Middle Ages, knowledge was predominantly codified by church scholars in line with the clerical principles (Kister 1981: 6). A new age of encyclopedia-making was initiated by the development of modern science, the advent of printing technology and a growing level of universal education. An encyclopedia started to be used as a source of information on novel technological developments presented in an easily accessible and condensed format (Kister *ibid.*, Katz 1998). The needs of the times and the advancement of knowledge in the 20th century have also exerted an impact on the genre, contributing to the versatility of encyclopedic forms and raising the standards of quality.

Today, due to the spread of the internet and digitalization, the genre undergoes further transformations. The internet itself is considered the largest reference service so far constructed, an encyclopedic compendium of knowledge, encompassing all kinds of data from every corner of the world. It has been even noticed that the advent of the web has initiated the process of gradual “opting out” of print reference works, due to the ease and convenience of usage of such works online (Castels 2002, Eco 2002). Moreover, a quick overview of encyclopedias available on the web shows that they represent a variety of forms, reflecting interesting properties of format and style, as compared to their print counterparts.

It is also due to the internet, and specifically the development of Web 2.0¹ technologies, that a close realization of Wells’s and Bush’s visions of the perfect encyclopedia has recently come into existence in the form of Wikipedia – the free online encyclopedia. Jimmy Wales, the founder of Wikipedia, had a dream similar to Wells’s, which was “to create and distribute a multilingual free encyclopedia of the highest possible quality to every single person on the planet in their own language” (www.en.wikipedia.org). We might say that his dream is slowly coming true. Wikipedia is a collaborative authoring project written by volunteers, striving to constitute a free repository of encyclopedic knowledge. It exists in over two hundred language versions, and has become the most popular encyclopedia on the internet. The free encyclopedia is currently treated as a source of information, data, facts and learning materials. As Viegas *et al.* (2007: 1) point out, “Wikipedia, the online encyclopedia, has gone from being a Web curiosity to becoming a powerful source of information for both the online and offline worlds.”

Still, Wikipedia constitutes an extraordinary and controversial structure. Voices evaluating the project range from the extremely positive, which praise the capacity of Wikipedia to amass large quantities of information, to those doubting the trustworthiness of such an open-ended, user-created reference web site. Moreover, the genological status of Wikipedia remains ambiguous. On the one hand, the free encyclopedia represents the transfer of the genre of encyclopedia to a new medium. At the same

¹ The term Web 2.0 has been introduced to encompass new phenomena on the internet. In contrast to Web 1.0, Web 2.0 projects put emphasis on user-generated content, the growth of social networks, user-interaction and multimedia.

time, though, it constitutes a product of the web, a unique digital genre, compiled in a distinctive and special context.

The emergence of Wikipedia and the controversies around the project raise the need for a discussion concerning the problem of genre evolution influenced by the medium of the internet. The evolution of genres, specifically, is closely associated with technological change and the advent of new information technologies (Fairclough 2005). Such a relation is clearly illustrated by the growth of the web which has been accompanied by a simultaneous increase in the number of different genres.

To date, however, no substantial research has been devoted to the evolution of the genre of encyclopedia, as conditioned by the web. It is not certain how and to what extent the medium influences such an evolution, and whether it instigates any modifications affecting the generic structure of this text-type. As far as Wikipedia is concerned, despite its widespread popularity, it still represents a largely underresearched area. The studies devoted to this encyclopedia conducted so far have mainly focused on the technological aspects of the system (Desilets *et al.* 2005, Capocci *et al.* 2006, Viegas *et al.* 2007) or on the specificity of the community of Wikipedians (Bryant *et al.* 2005, Holloway *et al.* 2005). There is insufficient research, however, into the genre properties of the free encyclopedia.

The current study is designed to fill that niche and address these questions. We assume that the medium exerts a significant impact on this particular genre, producing notable modifications of individual components of the generic structure. We also claim, with regard to Wikipedia, that the new context in which the free encyclopedia is composed will lead to the emergence of novel characteristics of function, content and form, distinct from those of standard encyclopedias. Further, we posit that the free encyclopedia will exhibit novel discourse properties reflecting the influence of the different environment in which it is created.

Specifically, the research project reported in this thesis has the following goals:

- to investigate the generic properties of encyclopedias available online with the aim of analysing the main types of encyclopedic web pages;
- to analyse the generic structure of Wikipedia, with the purpose of determining the degree of its equivalence with paper antecedents of the genre and delimiting the main areas of genre modification;
- to examine the discourse properties of Wikipedia articles in order to indicate how the style of an encyclopedia has evolved due to the influence of the context and how this evolution has affected the communicative purpose of the genre.

The analysis aims at a description of the main tendencies and general processes in genre evolution triggered by the medium of the web. It concentrates on the description of the genological status of articles in Wikipedia, their formal structure, stylistic features and mechanisms which govern the process of collaborative creation of this source of knowledge. In order to verify the degree of influence of the medium on the genre, the traditional model of an encyclopedia is compared with today's practice as exemplified by Wikipedia authors. Moreover, the study will allow an assessment of the degree of genre conventionalization and schematization of textual structures ascribed to an encyclopedia.

Materials and Methods

As far as the material for the study is concerned, the analysis of online encyclopedias is based on the investigation of 78 of the most popular encyclopedic sites currently available on the web. The material for the analysis of Wikipedia discourse comprises a collection of all articles beginning with letter *U*² available on Wikipedia, i.e. 968 entries.³ Such a choice enabled the inclusion into the material of articles of various length and subject matters. The articles were collected between December 2006 and December 2007.⁴

The first part of the description concentrates on a genological analysis of web pages. In this part, the analytical methods proposed by traditional genology are modified and updated to include medium specific properties of the genres appearing on the world wide web (i.e. page design, components and functionality).

The second part presents the description of the Wikipedia discourse, involving the evaluation of the dominant discourse properties of the entries, the most common sentence structures and properties of the lexicon. The aim of qualitative observations is to illustrate the nature of the information and the organization thereof.

² The letter was arbitrarily chosen from the alphabetical index of the free encyclopedia.

³ The few articles in the thesis whose headword is marked by a different letter constitute results of an automatic redirection made by the system from the original headword accessed in the index within the articles beginning with *U*.

⁴ It needs to be noted that Wikipedia is continuously evolving, with new articles being added on a daily basis and existing ones being continually modified. Therefore, the same collection of articles today might be considerably larger and somewhat different.

1 . Genology as a study

1.1. Schools of genre analysis

Genre research has its origins in the area of literature and rhetorics. For a long time, genre studies were concerned with structures and classes of literary texts. The term “genre” was usually referred to as a categorization according to form and topic (e.g. a tragedy, a comedy, an epic). In rhetoric, genres were determined on the basis of such characteristics as form, subject, situation and audience (Yates and Orlikowski 1992: 300). Today, genology encompasses a vast number of areas indeed, whereas the concept of genre is used in a number of disciplines, e.g. folklore, rhetoric, literature and art.

As Gajda (1993: 247) observes, two main periods of genre study may be outlined, a universalist stadium, ranging from antiquity until the 19th century, and a scientific stadium, from the turn of the 19th and 20th centuries onward. In the first period of genre study in the era of Plato and Aristotle, the main terms were constructed based on general philosophical and logical assumptions. This stage was largely normative in character, as genres were treated as a collection of strict and intransgressible rules defining the topic, composition and style of texts. Texts were analysed with reference to an ideal genre, constituting a model of structure and quality (Gajda *ibid*).

In the second period, the scope of genre description, its aims and methods gradually changed towards a more descriptive approach, focusing not only on textual regularities, but more importantly on the social and cultural aspects of genre use. Thus, genre was no longer treated as a static, classifying concept, perceived only in terms of a set of formal criteria, but rather with reference to context and purpose. As Freedman and Medway (1994: 2) argue, in the modern approach the focus is placed on the means to “unpack the complex social, cultural, institutional and disciplinary factors at play in the production of specific kinds of writing.” Genological studies nowadays aim at creating a complex, all-encompassing concept of genres, presenting their historical development, changeability, classification and the relations among them (Gajda *ibid*).

Within the field of linguistics, three main approaches towards genre analysis have developed, i.e. systemic-functional, applied linguistics and new rhetorical (Łyda 2007). What the three approaches have in common is the emphasis put on the purposefulness of the genre and its social meaning. What differentiates them are research techniques and the emphasis put on the static nature of genre (systemic-functional) versus on a dynamic nature of genre (new rhetorical).

1.2. Genre – characteristics of the concept

Within the present analysis we treat genre as a concept describing widely recognized types of communication structured according to substantive, compositional and linguistic conventions (Yates and Orlikowski 1992), directed at the achievement of a specific aim. The forces implicit in the structuring of genres encompass social, individual, historical as well as technical determinants (Erickson 1999: 2), which influence and regulate the content and linguistic form of a genre.

On the basis of this definition and the aforementioned theories, the basic characteristics defining a genre may be derived. Thus, a genre is typically described as:

1. A pattern of communication – meaning that a genre comprises not merely the particular pattern of a text, but is associated with the communicative situation, encompassing also the relation between the writer and reader or the speaker and audience. Moreover, the pattern refers to “genre” encompassing apart from single occurrences of events, also genre repertoires and hierarchies (Yates and Orlikowski 1992).
2. Situated – signifying that a genre occurs as a response to a particular recurrent situation and is embedded in the communicative event (Berkenkotter and Huckin 1995).
3. Dynamic – meaning that genres cannot be understood as static forms, but are subject to evolution due to different social, technological or linguistic factors (*ibid.*).

What the afore-described theories underline is the sharing or the recurrence of certain conventional features of genre construction. As Swales (1990) observes, a genre becomes recognizable only after it develops into a standardized configuration of specific elements. This means that apart from the recurrence of contextual features, a genre is characterized by the regularities of structural organization. The most important aspects determining a genre entail similarities in:

1. Form – encompassing repeated patterns in the structure of the genre, its individual components, discourse and compositional properties (format, structural layout, style-sheet) as well as medium of communication.
2. Content – (or substance in Yates and Orlikowski’s definition) referring to the topics covered by the genre.
3. Function – referring to the purpose and role a genre fulfils within a given community, a reason for the communication. The purpose of a genre constituted for a long time the central notion in genre identification and classification (Swales 1990). In reality, however, texts may fulfil many implicit purposes behind the one

conventionally associated with a specific genre. This observation led to the rise of many voices downplaying the role of purpose in genre characterization (cf. Fairclough 1995, Askehave and Swales 2001, Swales 2004). Still, it remains a vital component of genre description, being helpful in the differentiation between genres and sub-classes of genres as well as in the identification of the evolution of a genre.

Such a collection of determinants encompasses all the major aspects of genological description, i.e. the pragmatic, structural, cognitive as well as stylistic facets.

What constitutes the central aspect of a genre is that it is recognizable and standardized, being based on a set of conventions shared by the members of a community. As Bhatia (1997: 360) points out, irrespective of the way we refer to genres, e.g. as “a typification of rhetorical action”, as “staged goal oriented processes”, or as “shared communicative processes”, genres constitute products of the knowledge of conventions. Genre conventions give individual genres their identity or what Bhatia calls “integrity”, understood in terms of “a typical constellation of form-function correlations representing a specific communicative construct realizing a specific communicative purpose of the genre in question” (Bhatia 2004: 123). Genre conventions, thus, as the scholar observes, determine the choice of the genre to meet the conditions of a particular context and influence the use of particular linguistic means, that is lexical, grammatical and discourse means, appropriate to a specific situation.

It needs to be noted, however, that the conventions governing specific genres differ in their fixity and flexibility, which exerts an influence on the form of genre norms. Some genres need to realize an obligatory model, other forms are conventionalized only to an extent, whereas other genres are characterized by considerable flexibility in the realization of the generic conventions (Dobrzyńska 1992: 37). The flexibility of genre conventions also signifies that a particular instance of a genre does not need to replicate the model, but can modify it in different respects (Gajda 1993: 254). As Łyda (2007) observes, though genres exert considerable influence on the shape of an interaction, they do not determine it completely. Transformations and generic manoeuvres enable “continuous modifications, allowable variations and the subsequent evolution of genres” (*ibid.*: 31). A single genre can differ with respect to the historical period, culture or communicative setting, since users take into consideration the context and conventionally draw on other texts they are familiar with (Hyland 2002: 120).

1.3. Genre evolution

As emphasized above, genre constitutes an open, dynamic and evolving structure, adapting itself to changing social contexts, contemporary needs and purposes. As Mikhail Bakhtin (1986: 80) claims “genres are much more changeable, flexible and

plastic than language forms are.”⁵ Similarly, Berkenkotter and Huckin (1995: 6) state that “genres are inherently dynamic rhetorical structures that can be manipulated according to conditions of use, and that genre knowledge is therefore best conceptualized as a form of situated cognition embedded in disciplinary cultures.”

An interesting discussion of genre evolution was put forward by Yates and Orlikowski (1992), describing the development of the memo genre from the mid-19th century to the present. In their article “Genres of Organizational Communication” (1992) the scholars refer to the so-called Structuration Theory proposed by Giddens (1984). According to this theory, grounded in sociology, social institutions are produced and transformed owing to a community’s application of social rules. By means of these rules, individuals “reaffirm or modify the social institutions in an ongoing, recursive interaction” (Yates and Orlikowski 1992: 302). Drawing on Giddens’s theory, Yates and Orlikowski argue that the use of genres is governed by rules which associate chosen properties of form or substance with a particular situation. This means that a genre is situated in a stream of social practices that shape it and at the same time are reshaped by it. In using a genre participants identify genre rules from their experience and select a genre appropriate for the situation. During this process, individuals usually reproduce a genre but, as the scholars observe, they may also elaborate or replace it in order to adapt to a change of context, which results in a change of the genre.

The dynamism of genre, its susceptibility to evolution and transformation under various factors influences the perception of genre conventions. The identification of the determinants modifying these conventions constitutes a vital aspect of the study of genre evolution. In different historical contexts different factors exert impact on the development of a genre. Nowadays, as Bhatia (1997: 635) emphasizes, the changes in genres most often constitute the consequence of the necessity to be innovative in communication, of the influence of multi-disciplinary contexts and, what seems to be the most significant in the following study, of the use of multi-media and information technologies.

1.4. Discourse and genre competence

The discussion of discourse competence and the awareness of genre conventions and language practices seems useful in this context. The realization of the genre of an encyclopedic entry requires appropriate language skills and a degree of awareness of the communicative situation.

⁵ It needs to be mentioned that Bakhtin’s work on genres was written as early as in 1950s, with its fragments published first in 1978 in *Literaturnaja Uczoba*, 1, 200-219.

1.4.1 Discourse competence

Discourse competence refers to the ability to understand and construct texts of different genres (Duszak 1998: 252). What indicates discourse competence, as Duszak claims, is “the ability to produce texts which naturally fit into the cultural and linguistic conditions of a given communicative system generally and into the realized communicative act specifically” (*ibid.*: 253).

Generally, discourse competence concentrates on the text, on stylistic aspects, and the degree of text appropriateness. It encompasses also the awareness of the audience, the understanding of the effect which a particular text produced on the receivers (*ibid.*). Such awareness is fundamental in the production of texts such as an encyclopedia entry, where the precision, intelligibility and clarity of expression are particularly important.

In the light of the following study, discourse competence needs to be expanded so as to include successful operation in the context of the web and the specificity of mediated communication. As such, it needs to cover the ability of hypertext production, i.e. a text which is supposed to function within a larger network of interconnected texts. This involves the ability of creating meaningful hyperlinks from the main body of the text as well as the ability of a successful combination of different modes of representation. As Kenning (2006) observes, in this environment it is necessary to utilize both modes (textual and visual), as they may support one another, providing the reader with complementary information.

1.4.2 Generic competence

For a successful use of specific genres, it is vital to understand the afore-described conventions of genres as well as the practices of particular discourse communities, as these conventions, together with the social customs, make up the so-called “genre knowledge” or “generic competence” (Bhatia 2004: 145).

Genre knowledge constitutes a complex and dichotomous structure (Johns 1997: 21), being both cognitive (based on prior knowledge and experience) and social (shared with other members of the community). Genre knowledge, as Johns observes, provides a shortcut in the production and processing of texts. It involves the knowledge of the basic components of a text, such as the appropriate subject matter, tone and approach, together with the usual layout and organization.

Genre conventions give rise to the development of certain expectations users have towards specific text-types. These expectations, as Duszak (1998: 262) observes, are shaped by the community’s experiences in the usage of genres. It is with regard to these expectations that users evaluate texts as acceptable or improper in a particular communicative situation. Similarly, Bugajski (1993: 91) emphasizes that genre conventions delineate the limits of free usage of language means, the transgression of which causes violation of the rules which govern the communicative event and violation of the basic functions of the language.

The knowledge of conventions is important not only for the use of genre, but constitutes the basis for a modification of a genre, which has already been pointed out above. As Bhatia (2004: 208) emphasizes that “the knowledge of generic conventions should be best viewed as a resource to exploit generic conventions to respond appropriately to the requirements of professional practices rather than a blueprint for replication.” The fact that the knowledge of conventions enhances an individual’s performance within a genre is confirmed in the words of Bakhtin (1986: 80), who states that “genres must be fully mastered in order to be manipulated freely.”

Genre competence does not only concern the knowledge of contemporary genres and ways of communication, but also the knowledge of genres of the past. This seems particularly significant nowadays, in the era of the internet, where new genres are constantly emerging. Genre awareness, as Witosz (2005) observes, lets us perceive genological chains, mutual interdependencies between genres and rules of genre development.

As Gajda (1993: 252) emphasizes, genres represents an array of means which help us understand reality. They constitute a type of conventionalization of textual and spoken forms as determined by social, cultural and situational factors. The influence of these factors causes constant movements within genres, their evolution, change and the emergence of new structures.

2. Genres on the web

2.1. Genre theory on the web

The advent of the web has initiated far reaching transformations in the area of genology. The development of the World Wide Web has been accompanied by a dynamic growth in the number of various text types appearing on the web. Traditionally, the course of genre evolution and modification was slow, with new forms emerging gradually (Witosz 2005). What we are observing nowadays are dynamic and intense transformations of genre forms instigated by an array of new technologies.

Genre theory was introduced to the research of World Wide Web by Yates and Orlikowski (1992). The concept has since then been frequently exploited in the analyses of web pages and digital documents. Scholars underline the usefulness of the application of the genre concept in the study of the web and emphasize that genre studies can enhance our comprehension of the evolution of text-types and communities inspired by technology (Boese 2005). Crowston and Williams (1997) consider the web an excellent place for the study of new types of genres due to its architecture, devoid of a central management and the enforcement of genres, which enables freedom of expression and the possibility of experimentation. The scholars also underline the fact that the openness of the medium attracts many different communities which have experiences with different genres and use the internet for various purposes, which contributes to genre change. Similarly, Bauman (1999) stresses the importance of applying genre theory when studying the web, as, in her opinion, it encourages a better understanding of a text as a means of communication. She claims that web texts can be perceived not as equivalents of print forms, but as individual texts serving specific social functions which are valued by a particular discourse community.

2.2. The concept of web genre

The emergence and basic characteristics of the genre of digital documents were first described by Shepherd and Watters (1998), who regarded the World Wide Web as “a powerful trigger” (*ibid.*:1) of new text-types which they termed “cybergenre”.¹

As web genre development has been activated by the new medium, the focus of genre analysis has shifted towards the analysis of the capabilities of the web and the influence they exert on the evolution of genre characteristics. The analysis of web sites has concentrated on the technical factors and the properties of computer-mediated communication which result from the introduction of the new technology. These properties comprise hypertextuality, i.e. the interconnection between pages through the use of hyperlinks, interactivity and availability of multimedia (cf. Lanham 1993, Bardini 1997, Warschauer 1999, Jensen 1998, Bucy 2004).

The above-mentioned properties of the web have been brought together under the term “functionality”, introduced by Shepherd and Watters (1999) to encompass all the capabilities afforded by the new medium. The scholars argue that users approaching instances of web genre have certain expectations concerning not only the structure or the style, but more importantly, “they also have expectations with respect to functionality, i.e. how to interact with the genre and what to expect from it” (*ibid.*: 1). The scholars illustrate these expectations with an example of users’ interaction with a search engine – users input an item they wish to search and expect a list of website addresses relevant to their query.

Thus, the major difference between the characterization of traditional genre and web genre is the introduction of yet another element to the description, i.e. the functionality attribute. This means that while genres, as illustrated in the previous chapter, were traditionally characterized by the similarities in function, content and form, the description of web genre should be extended to include the aspect of functionality, referring to the capabilities afforded by the new medium.

Shepherd and Watters (*ibid.*) observe that it is the functionality afforded by the new medium that drives the evolution of web genre. It not only generates changes in traditional genres occurring in other media, but also causes the creation of new types of genres which have never existed in other environments (Erickson 2000). The majority of the studies devoted to the identification and classification of the genres on the web include this attribute as an integral characteristic, apart from the characterization of other components of a genre (cf. Erickson 1999, Crowston and Williams 2000, Herring *et al.* 2004).

¹ Though the term “cybergenre” was powerful at the time, it is now more commonly being replaced by “web genre”.

2.3. Genres on the web – taxonomy

The web constitutes a huge repository of texts exhibiting diverse forms. So far, a few groups of scholars have undertaken the investigation into the types of web pages.² The following table presents the results of the studies focusing on the extraction of the most common web genres:³

Research groups	List of web pages
Crowston and Williams (1997)	homepage, catalogue, archive, book, box store, chronicle, column, directory, e-mail, directory listing, essay, faq, guide, geneology, pamphlet, source code, game, brochure
Dewe et al. (1998)	informal, private, public, commercial homepage, forms, journalistic reports, faq, link collection, discussions, error messages
Shepherd and Watters (1999)	personal homepage, corporate homepage, brochure, resource, catalogue, search engine, game
Roussinov et al. (2001)	homepage, order forms, articles, news bulletin, special topics, publications, product information, product lists, glossary, link lists
Boese (2005)	homepage, calendar, documentary, e-card, faq, news, newsgroup, contact information, advertisements
Santini (2007)	e-shop, homepage, search page, blog, hotlist, e-mail, newsletter, online tutorial, splash screen

Table 1. List of the most common types of web pages

The first meaningful observations as far as the development and nature of web genres are concerned were provided by Crowston and Williams (1997) and Shepherd and Watters (1998). As the scholars noticed, three main groups of genres are domi-

² Recently also computer scientists have undertaken the investigation into the means of automatic genre identification. Such identification is considered particularly useful as far as engine searches are concerned, since the recognition of a particular genre could, as the scholars observe, significantly facilitate fast and convenient information retrieval. The most important research in this field comprises the studies of Karlgren and Cutting (1994), Kessler et al. (1997), Dewdney et al. (2001), Shepherd et al. (2004).

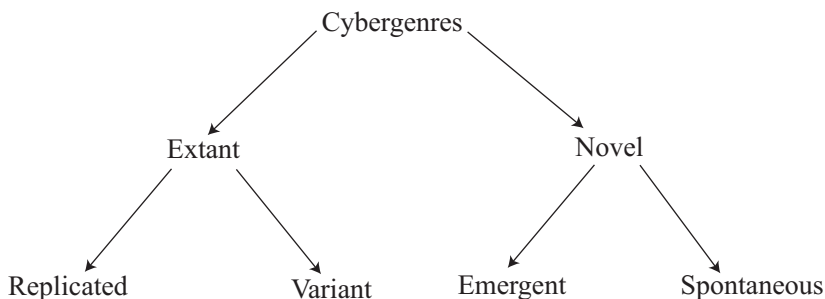
³ A few scholars concentrated on the analysis of individual web genres, e.g.: Bergenquist and Ljungberd (1999) studied the use of organizational document genres at work. Analysing the use of internal e-mail, they observed that many messages had the form of informal conversations. They concluded that the concept of genre is useful for analysing how communication establishes organization. Cukier et al. (2006) tried to apply the concept of genre to the analysis of the set of digital documents known as “spam”. They concluded with the observation that spam is not a single genre, but is represented by many genres, mainly adaptations of printed forms to the internet, e.g. business cards, pamphlets, advertisements or order forms. They also observed that spam exploited genre by conforming to accepted forms, while simultaneously breaking the norms.

nant on the web: genres inherited from other media, those that constitute mutations of other genres and genres which exhibit novel characteristics.

Crowston and Williams' (1997) study involved the classification of a set of randomly selected web pages in order to enumerate the types of genres in use. The taxonomy proposed by the scholars includes four types of genres: reproduced genres, constituting 61% of all the examined genres, adapted genres, encompassing 28% of the pages, novel genres (6%) and a group of unclassified web pages (5%). As visible from the percentages, most genres identified by Crowston and Williams constitute the group of reproduced genres, borrowed from other media. The scholars also identified a reasonably large group of adapted genres, modified to suit the capability of the new medium. As the scholars observe, the novel and unclassified web pages illustrate the process of evolution of the medium. Novel genres (i.e. home page, hotlists, interactive pages) serve communicative purposes which are unique to the web, whereas unclassified pages are considered to be genres in evolution, in the process of formation.

Shepherd and Watters (1998) in their study placed more emphasis on the analysis of the degree of functionality offered by the sites. Based on differences between the sites in this respect, they propose the following taxonomy of web genres:

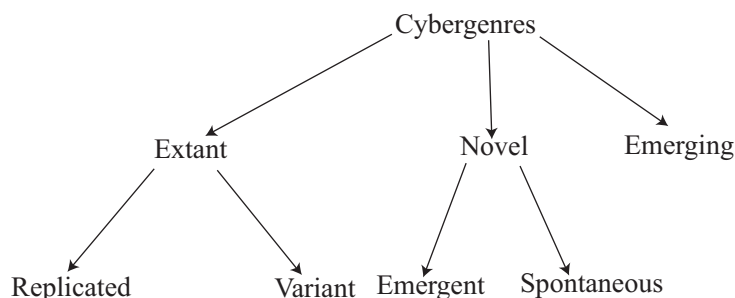
They divide cybergenres into two classes, extant and novel genres. Extant cybergenres are based on genres existing in other media, whereas novel genres are fully dependent on the digital media. The two classes are further divided to include more specific sub-types of genres. Replicated genres constitute equivalents to Crowston and Williams' reproduced genres, while variant subgenres are analogous to adapted genres in the latter taxonomy. According to the scholars, the taxonomy can be regarded as the stages of evolution of the genre – due to the functionality afforded by the medium replicated genres evolve into variant forms until novel genres emerge that are significantly different from the original genres.



The main difference between the two taxonomies concerns unclassified web pages and novel genres. Shepherd and Watters (1998), unlike Crowston and Williams (1997), do not include unclassified web pages in their taxonomy. Crowston and Williams state that new genres observed on the web are in most instances derived from the already existing genres and have only been adapted to the conditions of the new

medium, whereas Shepherd and Watters propose the differentiation between such genres, termed as emergent, and spontaneous genres, which, in contrast to emergent forms, do not have any previous counterparts in other media.

We agree with Santini (2007) that the two taxonomies complement each other. Judging by the present situation on the web, we are facing both spontaneous genres as well as pages which defy classification, i.e. emerging genres, and their presence also needs to be recognized. Thus, combining the two taxonomies offers a wider and all-encompassing view on the types of genres found on the web:



The respective groups of web genres can be described as follows (Crowston and Williams (1997), Shepherd and Watters (1998), Santini (2007):

a) Reproduced genres

Reproduced or replicated web genres represent genres transferred to the web from other media. Such web pages closely reproduce the generic features of their counterparts existing in other media. The level of functionality added by the new medium is scarce and does not impinge on the structure and content of the genre. As examples of replicated genres we may include scientific journals, official forms, dictionaries, advertisements and scholarly papers published on the web.

b) Adapted genres

Adapted or variant web genres, in contrast to reproduced types, exhibit a degree of differences in structural and pragmatic aspects as compared with their counterparts in other media. What is particularly important is the higher degree of functionality added to the genre. Instances of adapted genres are e-zines, e-books. It needs to be noted, though, that at times it may be difficult to evaluate the level of functionality needed to mark the boundary line between replicated and adapted genres.

c) Emergent genres

Emergent genres represent new web genres whose conventions and rules of usage can be delineated and have been acknowledged by their authors and users. Emergent genres manifest a considerable departure from the generic features characterizing their original counterparts, which exerts an influence on all the components of the

genre model and leads to the development of novel forms. Emergent genres are characterized by a significantly developed functionality attribute, which makes them dependent on the new medium. Emergent genres are abundant and include web pages such as blogs, chats and e-mails.

d) Spontaneous genres

Spontaneous genres, in contrast to the other types of web genres, are based completely on the new medium. Such genres do not have any counterparts in other media and have emerged together with the development of the new medium. They are shaped by the conditions and rules governing the functioning of the web. Spontaneous genres include: FAQs, homepages, search engines, error pages, discussion groups, newsletters and portals.

e) Emerging genres

In contrast to emergent genres, acknowledged by the users, emerging genres are in a stage of development and lack clear genre conventions, and thus cannot be included into any of the listed genres. The emergence of a new genre, according to Santini (2007: 5), may be hypothesized on the basis of the existence of a recurrent textual pattern which does not have any conventionalized name.

As Santini (*ibid.*) observes, the class of emerging genres deserves attention since it embodies a transitional phase characterizing many pages available on the web and indicates the path of web genre development. As she rightly emphasizes, many of the nowadays fully conventionalised emergent genres (e.g. blogs) were at their initial phases considered emerging, unclassified web sites, until they became acknowledged by the community of users.

2.4. Genre hybridization and individualization

What should be also noted is the growing tendency towards genre hybridization and genre individualization, which is evident on the web (Santini 2007: 2). Genre syncretism, as Santini (*ibid.*) claims, results from the fact that web genres have the form of web pages, which constitute complex structures, involving various communicative purposes and functions. As the scholar observes, a web page may be treated as a sort of a container of multiple texts “where not all the elements necessarily belong together, but they all contribute to form a whole, even without any linear progression” (*ibid.*). Such complex pages are very popular mainly owing to the use of hypertext, which allows the introduction of different generic structures and forms within one page. Multi-genre web sites comprise for instance homepages, online newspapers and portals, which include in their structure an range of different genres (cf. Grzenia 2006).

The architecture of the medium, and the above-mentioned freedom of variation can instigate yet a different process, namely that of genre individualization, which refers to authorial experimentation leading to the production of novel forms (Santini 2007: 7). On the web, as the scholar observes, constraints are considerably lesser than

in other media, which gives the authors of web pages more freedom to alter traditional genre patterns.

2.5. Web genre evolution

As observed by scholars (cf. Breure 2001, Boese 2005, Santini 2007), the process of genre evolution on the web, unlike in other media, is much more visible and dynamic due to the fluidity of the medium plus a greater degree of autonomy of the users, which stimulates inventiveness. As argued by Yates and Sumner (1997) and by Bauman (1999), novel generic forms arise as a result of the influence of the new medium and in response to the changes of users' needs.

To explain the factors underlying the evolution of web genres, Rehm (2006) proposes a differentiation between four phases, i.e. genre production, modification, change and reception. He assumes that authors of web pages incorporate elements of already existing web sites, which leads to the emergence of specific conventions governing the use of these genres. If authors break the conventions, the process of change is initiated. Such an approach to the evolution of web genres reminds us of the conception proposed by Yates and Orlikowski (1992), who drew an analogous circle of genre evolution.

To sum up, we may state that the speed and intensification of technological as well as social and cultural changes causes a considerable complexity of generic forms. Digital genres assume different and more dynamically changing forms due to the progress of information sciences, but also owing to the activity and creativity of the users of the web.

3. Encyclopedia as a genre

3.1. Encyclopedia as a genre of information

Encyclopedias have a long tradition of use, constituting a generally familiar genre. As Louis Shores, editor of *Collier's Encyclopaedia*, put it, "the importance of encyclopedias lies in their ability to help people see that knowledge has unity, and thus help to achieve generalization in a world of over-specialization" (after Collison 1964: 213). Similarly, Kister (1981: 4), an encyclopedia critic, argues that "whatever their imperfections, general encyclopedias can serve as a gateway to understanding the most profound or intricate knowledge human beings have yet produced." Yet, as Olkiewicz (1988) observes, there is usually something more to an encyclopedia. These compendia of knowledge rarely constitute only impassionate and strictly objective registers of information. As the scholar emphasizes, each encyclopedia "emanates the spirit of its time, follows a particular system of values, as it mirrors the views and mentality of its creators, but also of the receivers to whom it is addressed" (*ibid.*: 8).

Encyclopedias belong to the large group of genres of information, the collection comprising all sorts of reference materials, i.e. sources consulted to discover specific facts or information.

Due to a large internal diversification of reference materials, their typology may be performed according to different criteria, such as scope or function (Marszałek 1986, *Encyclopedia Britannica* (EB) 1990, *Wielka Encyklopedia Powszechna* (WEP) 2002). The major criterion, nevertheless, involves the type of information included in the work, i.e.:

Factographic	→	encyclopedias, lexicons, yearbooks, almanacs
Lexicological	→	dictionaries, glossaries
Geographic	→	maps
Statistic	→	statistical compendia
Lists	→	bibliographies, outlines, catalogues, directories (Marszałek <i>ibid.</i> : 140).

So far, few studies have concentrated on the genre of encyclopedia and few have provided a full and complete description of the components of this genre, possibly due to a large internal differentiation of the sub-types of encyclopedias, their varying content and structure.

An encyclopedia is by definition a “compendium of knowledge, either general, attempting to cover all fields, or specialized, aiming to be comprehensive in a particular field”,¹ mainly presented in an either alphabetical or systematic order. Further definitions provide complementary explanations stating that an encyclopedia “gives authoritative information on the covered field of knowledge”² or “background information and an overview of a particular subject, person, place, or event. An encyclopedia summarizes and synthesizes information from a variety of sources and often includes a selective bibliography of authoritative books and articles on a topic.”³

3.2. Genre characteristics of an encyclopedia

The enumeration and description of all types of encyclopedias is virtually impossible, considering their contemporary state of advancement. Yet, though there is no standard, accepted method of classification of encyclopedic works, they are usually grouped according to certain major criteria:

30

Scope	→	universal – encompassing information from all areas of knowledge; specialized – concentrating on one discipline or a large concept.
Size	→	large, multi-volume containing extensive information; small, one-volume with a concise compendium of information.
Classification	→	alphabetical; systematic.
Author	→	collective – where the articles are written by different authors; individual – being the work of one author (cf. Marszałek 1986, EB 1990).

Though there does not exist a standard procedure guiding the process of compiling an encyclopedia, a number of principles governing the production of an encyclopedia which delineate the aspects of function, form and content of the genre have been ac-

¹ Columbia Encyclopedia <http://www.bartleby.com/65/>

² Online Dictionary for Library and Information Science (ODLIS) http://lu.com/odlis/odlis_e.cfm

³ Library Terms and Definitions <http://www.lib.uoguelph.ca/help/research/terms.htm>

cepted throughout the evolution of the genre. This has led to a relative “ritualization” (Fairclough 2005: 72) of the genre, in that its generic structure is characterized by a more or less predictable set of components, which encompass the following:

- Function: – a source of information providing only necessary data generally in a concise, clear, objective and communicative way;
- Content: – articles of any substance prepared by specialists;
– thematic scope either universal or specialized;
– adjuncts:
– illustrative material (optional);
– bibliographies appended to longer articles;
– an analytical index of people, places and minor subjects;
– addition of supplements to bring the main work up to date;
- Form: – segment structure made up of individual textual units;
– systematic or alphabetic classification;
– system of direct and indirect references, which unite the entries into hierarchically ordered wholes;
– specific discourse properties.

31

Further characteristic properties of encyclopedic texts, referred to as a “Cinderella” or “colony” text-type⁴ (Hoey 2001: 73) involve the following:

- the meaning of the whole text is not derived from the sequence of the components;
- adjacent units do not form continuous prose;
- there is a framing context for the interpretation of the text; there is no single named writer;
- one component may be used without referring to the others; components can be reused, added, removed or altered;
- all components serve the same function (*ibid.*: 88).

The main properties of encyclopedic works, developed throughout the centuries, which define this genre of reference are:

- authority – achieved by the collaboration of experts on the preparation of the content;
- accuracy – achieved by the provision of reliable and valid information. Nevertheless, despite the care of the publishers errors of fact may occur;
- currency – achieved due to various systems of updating. However, due to lengthy publication processes and infrequent re-editions encyclopedias frequently fail to provide most up-to-date information (Olkiewicz 1988, EB 1990, WEP 2002).

⁴ “Cinderella” or “colony” texts involve texts which are homogeneous with respect to discourse characteristics, yet exhibit heterogeneity with regard to their appearance and use. Apart from encyclopedias, Hoey classifies within this group texts such as shopping lists, statutes, hymn books, newspapers, cookery books, journals and footnotes.

As the above listed principles concerning the composition of an encyclopedic work have not lost their validity and determine the genre characteristics of contemporary encyclopedias it would be relevant to discuss them in greater detail.

3.2.1. Function

To put it concisely, an encyclopedia is a storehouse of facts. As Grogan (1987: 2) describes it, “the encyclopedia takes all that we know apart into little pieces, and then arranges those pieces so that they can be found at a time.”

The aim that authors of encyclopedias had in mind when creating their works evolved throughout the ages, depending on the time and approach to knowledge and education. The Greeks saw the role of an encyclopedia as an enhancement of the process of thinking, the Romans perceived the issue more practically, as an encyclopedia was intended to help them become more competent and work more effectively, whereas mediaeval Christians treated knowledge as a means toward absolution (EB 1990: 258). Similarly, Zedler stated in his *Universal Lexicon* that “the purpose of the study is to combat atheism, and to prove the divine nature of things” (after McArthur 1986: 155). What all these approaches have in common is the conviction that an encyclopedia could contribute to the perfection of mankind (EB *ibid.*). Vincent of Beauvais, for instance, maintained that the ideas presented in his *Speculum maius* should be followed, as their goal was to improve the world. Francis Bacon, similarly, believed that his writings could contribute to the enhancement of the intellect of the ordinary man (Collison 1964). Samuel T. Coleridge saw the function of his *Encyclopaedia Metropolitana* to teach “Philosophy in union with Morals; and to sustain Morality by Revealed Religion” as well as “to convey the pure and unsophisticated knowledge of the past so as to aid the progress of the future” (Coleridge 1818, reprinted in Collison *ibid.*: 297).

Today, an encyclopedia is perceived rather as a basic source of reference, an immediate source of facts, a tool for finding general background information and explanatory material or definitions (Jacobson 1995: 433–4). Considering the contemporary approaches to knowledge, it may seem that what matters most is fast access to information as well as the possibility of finding data which are, as Grogan (1987: 3) puts it, either “too restricted in scope or too abstruse in content to justify a whole book.” Moreover, an encyclopedia can serve as a source of “pre-research” information (Jacobson *ibid.*: 434), or as a point of departure towards the discovery of further knowledge, which can be performed by following up the references given to the articles.

3.2.2. Content

3.2.2.1. Authorship

In the early period of encyclopedia-making, the works were compiled by individual authors, who, in the majority of cases, undertook the creation of these works of reference on their own (e.g. Pliny, Varro, St. Isidore).

Collaboration in the production of encyclopedias became common at the beginning of the 18th century (e.g. Harris's *Lexicon*).⁵ This approach was based on the conviction that it is simply unattainable for an individual to understand and present all possible knowledge (Yeo 2001). Such perspective may be observed in the words of d'Alembert "what man, then, could be so brash and so ignorant in understanding as to undertake single-handedly to treat all the sciences and all the arts?" (after Yeo *ibid.*: 79).

Nowadays, encyclopedias are in the main written by groups of experts, authorities in a given field of knowledge. In addition, the composition of an entry, selection of the information and overall structure of the encyclopedia are in general supervised by teams of editors.

The content of the entries may be either attributed or non-attributed to single authors. Most frequently, though, the entries are not signed by the writers. This anonymity of encyclopedias, as Hoey (2001) observes, influences the way they are perceived by users, who regard encyclopedias as arbiters of fact, offering a final court of appeal.

3.2.2.2. Scope

More importantly, encyclopedias differ in the thematic scope of their content, being either general or specialized. Due to the recently observed specialization of particular fields of science, the number of encyclopedias restricted to a specific subject has risen.

The delimitation of the boundary between the information belonging to universal or specialist areas is a matter of considerable importance as far as the scope of an encyclopedia is concerned. However, as Mielczarek (1972) observes, the interpretation whether particular terms are suitable either for a universal or a specialist encyclopedia requires a more profound understanding of the relation between common and specialized knowledge and vice versa, which frequently proves difficult to achieve.

The scope of an encyclopedia and its subject matter determine the potential audience of the work, involving either the general public, adults, specialists or children (Katz 1982).

3.2.2.3. Adjuncts

a) Illustrative material

Varro's *Disciplinarum*, edited around 50 B.C., is believed to be the first illustrated encyclopedic work. The work included a section on the great individuals of ancient Greece and Rome accompanied by around 700 portraits. Illustrations were also introduced to the work of St. Isidore *Originum seu etymologiarum libri XX*, and to a greater extent to his follower's Magentius Hrabanus Maurus *De universo*, which

⁵ The first noted example of a collaboratively produced encyclopedic work was *Rasa'ulu Ikhwan al. Safa's* produced by the Brothers of Purity, a religious and political party founded at Basra around 980. As Collison (1964: 41) notes, the aim of the party was the construction of a universal system of religious philosophy as well as the achievement of harmony between authority and reason. The compilation of 52 pamphlets written by five authors was issued with the aim of achieving the purification of knowledge and included information on arithmetic, the universe and the senses as well as on religion and spiritual life.

included 361 illustrations. It is also worth mentioning the *Hortus deliciarum* compiled by Abbess Herrad around 1170, the first encyclopedia compiled by a woman. With its nine thousand figures in 636 miniatures the work is considered as one of the best illuminated encyclopedias (Stockwell 2001). Nearly all the mediaeval and later encyclopedic works contained illustrations inscribed in the format of plates attached to the text. Also the first editions of the French *Encyclopedie* and the *Encyclopedia Britannica* are famous for their collection of beautiful illustrations including respectively 2,885 and 160 copperplates with explanatory texts. It is worth mentioning the Italian *Enciclopedia italiana di scienze, lettere ed arti* (1929–39), as it is widely acknowledged that the quality of the reproduction of the illustrations in this work reached an extraordinary level (Collison 1964: 207).

In modern times, illustrative material constitutes a standard component of an encyclopedia.

b) Bibliographic lists

Bibliographic material may help users obtain more data concerning the information given in an entry. To quote Collison (1964: 16): “in the modern world of interdependence between the various branches of knowledge, no article can be self-sufficient: it must be qualified, elaborated, heightened and illuminated by what is contained in the articles on related topics.”

The trend of introducing bibliographic records into encyclopedic works is believed to have been initiated by Guglielmo di Iacopo da Pastrengo in his work entitled *De originibus rerum libellus*, compiled around 1350, which contained an alphabetical bibliography of past writers (Collison 1964). From that time, lists of the most significant references and bibliographies have become a common component of an encyclopedic work. Encyclopedias differ in the manner of source listing, attaching them either to articles or arranging all the sources as a separate unit in the final volume of the work (Grogan 1987).

c) Indexes

Indexes constitute a relatively novel ingredient of the encyclopedic content. Though the first indexes had already started to appear in mediaeval times (e.g. Vives' *De disciplinis*), encyclopedias regularly started to include an index around the 1830s (Stockwell 2001). Among the English works, one of the first indexes to appear was that included in the 7th edition of the *Encyclopedia Britannica* (1827–42).

An index can be of great importance and assistance to users, as it greatly enhances the speed and efficiency of the process of consulting an encyclopedia. For an index to be useful, though, as Urbańczyk (1991) notes, it is important that it should refer the reader to a more or less exact part of the page, as the finding of specific data from among hundreds of words may present problems.

d) Supplements and updating

The importance of the currency of an encyclopedia started to be underlined around the 19th century. It was at that time that new systems of updating and supplementing

were designed (EB 1990). One of the most frequent methods is the publication of year books, presenting recent events and discoveries.

Still, earlier editions of encyclopedias, though they become out-dated, do not entirely lose their value. Grogan (1987) emphasizes the historical merits of previous editions and the account they can provide of the state of knowledge in a specific subject at the time the work was compiled.

3.2.3. Form

3.2.3.1. Density

Encyclopedias exhibit great variation as far as their size and scope are concerned. There are one-volume encyclopedias on offer (e.g. *Columbia*, *Petit Larousse*, *Brockhaus*). Yet, most frequently encyclopedic works number 20 to 30 volumes.

The cost of writing and printing technologies led to considerable constraints put upon the scope and the length of particular articles. As Pang (1998) observes, the need to save space requires careful planning and imposes the need for contractions, the presentation of entries as blocks of compressed texts and restrictions upon bibliographic references.

3.2.3.2. Article structure

The article constitutes the basic construction element of any encyclopedia. A prototypical encyclopedic article is recognized as a brief introduction to a particular subject written in a form of a compressed summary focusing mainly on the most important aspects of an issue (Stvilia *et al.* 2005). In contrast with the composites of other books, i.e. chapters, sub-chapters or paragraphs, an encyclopedic entry represents an individual unit which has a degree of autonomy (Marszałek 1986).

The length of articles exhibits considerable diversity, depending on the scope of the work and its degree of specialization. The structure of a complex article usually comprises the headword, definition, historical note, information about the contemporary state of knowledge and references.

3.2.3.3. Classification

The organization of an encyclopedia, i.e. the principles which guide the arrangement of content, constitutes one of the most important aspects of any encyclopedia.

Encyclopedic content was arranged either systematically (by broad themes or specific subjects) or alphabetically (Mielczarek 1972, McArthur 1986, Bolter 2001).⁶ Each of these methods was popular in different periods of the historical development of the genre.

⁶ Early encyclopedic works frequently resorted to different kinds of graphic illustrations representing the classification of knowledge. These encompassed representations in the form of tree diagrams showing transition from general to more specific concepts, or in the form of globe or map metaphors (Yeo 2001)

The systematic and associative categorization prevailed mainly during the initial years of the genre development, when an encyclopedia was seen as a work which was to be read as a whole. Systematic categorization was at that time perceived as the only satisfactory way of illustrating the unity of knowledge (Katz 1982).⁷ Alphabetical indexing gained popularity when the encyclopedia started to be treated mainly as a source of quick reference (McArthur 1986).⁸ One of the first encyclopedic works which introduced alphabetical ordering was *De proprietatibus rerum* written by Bartholomaeus Anglicus between 1220 and 1240, one of the most popular encyclopedias of the late Middle Ages (Collison 1964). The actual introduction of alphabetical arrangement is ascribed to the dictionaries of Robert Estienne (1499–1559) and Henri Estienne (152–1598) (Olkiewicz 1988).

As Bolter (2001) claims “the shift from hierarchical to alphabetic organization in encyclopedias was an acknowledgment that such systems as the seven liberal arts, which could be possessed by all educated readers, could no longer accommodate specialized knowledge in physics, anatomy, geography and mathematics” (*ibid.*: 85). A further advantage concerns the egalitarian character of alphabetic classification, since it avoids hierarchization of knowledge, reducing all the fields to the same level (*ibid.*). Moreover, such organization facilitates editorial work, allowing indefinite addition or deletion of information, without seriously impinging on the overall structure of the work (Stockwell 2001).

A serious disadvantage of alphabetical ordering, which is frequently emphasized (EB 1990, Stockwell 2001), is the perspective that it presents knowledge atomistically, thus separating the concepts which are thematically related. Contrary to alphabetic arrangement, systematic classifications treat concepts wholistically, illustrating their internal and external relations. Systematic approaches are thus suitable for readers who do not only seek concrete information, but a help in understanding complex issues from different fields (Stockwell *ibid.*).

Yet, there were encyclopedias which tried to unite alphabetical and systematic categorization. Both the English *Cyclopedia* and the French *Encyclopédie* presented diagrams illustrating the historical and logical relation of various subjects, though presented within the alphabetically arranged entries. One of the best examples of an attempt to unite alphabetical ordering with a philosophical vision is *The Encyclopaedia Metropolitana*. The title says: *a Universal Dictionary of Knowledge on an original plan, projected by the late Samuel Taylor Coleridge, comprising the twofold advantage of a philosophical and an alphabetical arrangement*. The authors of *Encyclopaedia Metropolitana* tried to exploit the ideals of an encyclopedia as proposed

⁷ For instance, Pliny’s famous *Natural History* was divided into entries on the stars and planets, then moving on to the geography of the Earth, humans, animals and minerals. The encyclopedia of St. Isidore presents yet a different arrangement of the content. The collection was divided into twenty books presenting information on e.g. the liberal arts, medicine, the church and heresis, the heavenly hierarchy, zoology and the man. This proves that no standard forms of categorization were available at that time.

⁸ Alphabetical arrangement was a late invention. It originated with the Greeks and their use of the “glossa” to describe difficult words on the margins of papyrus rolls (Katz 2003: 407), which evolved into a glossary, a word list.

by Coleridge in his *On Scientific Method*. For Coleridge, an encyclopedia was an educational tool, which should present knowledge to the readers appropriately, with the relations of things united not by “the accident of initial letters”, but by an idea. Coleridge, similarly to early encyclopedists, saw an encyclopedia as a work to be read from the first to the last page, with the text presented in a linear order (Collison 1964).

One of the most recent examples of a combination of alphabetical and systematic organization of knowledge is the 15th edition of the *Encyclopedia Britannica*. The encyclopedia is divided into three main parts, the *Micropaedia*, containing short, factual articles presented in alphabetical order, the *Macropaedia*, a collection of topically organized extensive and scholarly articles, and the *Propaedia*, which contains a classification of knowledge and serves as a topical guide to the articles in *Macropaedia* and shows the relationships between them. Such an arrangement of content combines a quick reference work with a topical set suitable for more demanding readers. Still, reviewers of encyclopedias (Kister 1981, Katz 1982) underline that such an organization is not particularly user-friendly.

Analogously, H.G. Wells maintained that an encyclopedia should not be a miscellany, but a clarification and a synthesis, while O. Neurath, an Austrian sociologist, proposed that an encyclopedia should be designed like an onion, where different layers of knowledge surround the heart, the foundation of unified science (EB 1990: 260, Stockwell 2001: 99).

3.2.3.4. Cross-references

The basic tenet underlying the structure of any encyclopedia is the accessibility and retrievability of the information enclosed in the work. The organization, as Jacobson (1995: 433) puts it, should “illuminate the relationships and intersections between innumerable areas of understanding.” One of the basic means of securing the interrelationships is cross-referencing.

A regular inclusion of systems of cross-references dates back to around 1410 with the work of Domenico Bandini *Fons memorabilium universi*, with interconnected sections on history and geography (Collison 1964: XIV). It is with this work that inter-relations between articles started to be marked.

A well-designed system of cross-references may facilitate the use of the encyclopedia by combining thematically related articles. Encyclopedias have developed a number of types of cross-referencing. The most common are arrows followed by the headword of the articles dealing with the subject in question (e.g. the *Encyclopédie* → Diderot), usage of the word “see” or capital letters.

3.2.3.5. Discourse properties

Historically, encyclopedic style evolved along the development of the genre. Early forms, especially those compiled by accomplished writers or distinguished personalities of the time, exhibited more individuality of style. As McHenry (2003) observes on the example of the 11th edition of the *Encyclopedia Britannica*, many entries of that time were written in a grandiose, exceptional style. With time, however,

the freedom of expression was gradually curbed by editorial requirements by the preference of standardized forms.

Nowadays, paper encyclopedias dedicate a lot of effort to guarantee a formatting regularity within the articles. The language of encyclopedia entries constitutes a significant uniting factor promoting the homogeneous nature of the work, and as such it is considered one of the distinctive properties of an encyclopedia (Marszałek 1986).

Moreover, an appropriate definition of specific headwords decides about the value of the work. Any definition in order to fulfil its function, from a stylistic point of view, should meet a number of requirements, of which the following are considered significant:

1. conciseness and sufficiency of information,
2. communicativeness and intelligibility,
3. objectivity, emotional neutrality, impersonality,
4. terminological precision, exactitude, clarity, unambiguousness,
5. logical consistency of the explanation,
6. use of defined technical terms, symbols and figures (cf. Gajda 1982, Marszałek 1986, EB 1990, Dąbrowska 1991, Wilkoń 2002, WEP 2002).

Conciseness, being closely associated with density and brevity, is a direct consequence of the limitations of space available to each article (WEP 2002). Space restrictions require as brief a presentation of the data as possible. Yet, though expressed concisely, the information enclosed within the definition should be adequate, which means that the definition should not be either excessive or limited, but sufficient to allow the comprehension of the concept discussed (Marszałek 1986).

Furthermore, an encyclopedic article should be presented explicitly, in an intelligible manner (Gajda 1990). Intelligibility should facilitate the use of an encyclopedia by a large group of readers, irrespective of their knowledge of the subject (Marszałek 1986, *MSN Encarta*). Intelligibility may be enhanced by proper structuring of the information i.e. a clear and coherent structure of the definition, which requires the primary information to be provided at the beginning of an article.

To be a relevant source of knowledge, the information comprised in the work should be reliable and objective. Objectivity, considered one of the most important features of encyclopedic texts (Grogan 1987), is often equated with matter-of-factness and accuracy (Dąbrowska 1991). A description will be accurate if it represents a particular fragment of reality precisely, in a manner free from any biases and prejudice by the author. Objectivity involves as well a fair presentation of controversial issues, with a description of opposing opinions on the subject (Grogan 1987).

Most major encyclopedias have faced accusations regarding bias and subjectivity (EB 1990). The history of the genre shows that encyclopedias frequently constituted an arena for the expression of authors' opinions or the cultural background of the country of origin. Many encyclopedias constructed until the 19th century fulfilled an ideological role, of which the most notable examples are the French *Encyclopédie*, or the early editions of the *Encyclopedia Britannica*, accused of prejudice and selectiveness. Though objectivity is difficult to achieve, modern encyclopedias increasingly

strive to be neutral and fair-minded, avoiding straightforward favouritism and a biased treatment of controversial subjects (Grogan *ibid.*).

The above-mentioned requirements correspond closely to Grice's Maxims of Manner, Quality, Quantity and Relevance.

Naturally, the properties of discourse may vary as they depend on the subject of the encyclopedia, its scope as well as the potential group of receivers at which it is targeted. The features described above are especially relevant as far as universal, general encyclopedias are concerned.

The afore described properties and components of the genre have evolved over many centuries, worked out to enhance the functional value of the work to suit the different needs of the users. Obviously, specific works differ in their construction, but the majority of the general encyclopedias can be observed to follow these conventions.

4. Encyclopedias on the web

4.1. Online reference works

As described above, the web has become an effective instrument for the creation of new genres, which involve also modern, electronic reference works. A selection of the encyclopedias available online have evolved from their print counterparts, whereas others function solely on the web.

The first online reference materials appeared on the web around 1992 (WEP 2002). As Buckland (2007) observes, the new environment has not changed the function of reference sources, but has created new tools for accomplishing their purposes. Reference sources still fulfil the role of a place where it is possible to seek and verify data as well as a source indicating the relations between different fields of knowledge (Buckland *ibid.*: 88). What has changed is their content and form, which is being continuously enhanced due to the ongoing development of technological affordances. It is claimed that the advent of computer technologies has revolutionized the way of creating reference sources, as it is possible to create large corpuses of data and configure different types of reference materials meeting the needs of various groups of users. Progress in the field of information storage, retrieval and distribution has proved indispensable also in editorial and revising processes (Piotrowski 2001).

However, online reference materials, in addition to such a wide range of options, offer also a range of new problems concerning their authorship, reliability and trust worthiness. Moreover, electronic technologies, as Bolter (2001) observes, have contributed to a change in our perception of knowledge. The scholar argues that in the mediaeval period the technology encouraged a stable organization of knowledge, whereas nowadays knowledge has become rather unstructured, assuming different forms and patterns according to the needs of the readers.

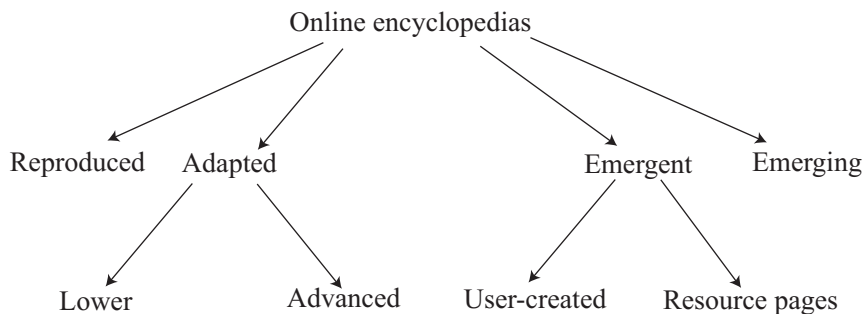
4.2. Typology of online encyclopedias

The studies on reference works conducted so far prove that any typology of such materials is to a large extent fluid (de Schryver 2003, Hartmann 2005). The same applies to a classification of electronic encyclopedias. The progress in computer technology has resulted in an abundance of electronic encyclopedias exhibiting great diversification. More and more highly innovative encyclopedias or reference sites are being developed solely for the web. Still, there is a lack of studies analysing the variety of encyclopedias available online. Therefore, the following classification should be treated as introductory to the subject and as a simplified representation of the myriad of encyclopedic sites. In the following study, we concentrate only on the encyclopedias available on the internet, omitting those issued on CD-ROMs or DVDs.

In order to identify how, and if so, to what extent, the transfer of the genre of encyclopedia to a different medium has affected its generic characteristics, we conducted an analysis of the most popular encyclopedias available on the internet. The collection of encyclopedias for analysis was extracted on the basis of a Google search for the “encyclopedia” query. The results of the search involved 113,000,000 web site addresses, of which the first 100 were evaluated. The final sample for analysis, after excluding repeated addresses, encompassed a collection of 78 encyclopedias.

The identification of genre characteristics and the production of a typology of online encyclopedias was performed according to the approach presented above, analysing the genre concepts of function, content, form and functionality. Content refers to the topics and materials offered by encyclopedias. Form refers to the design of the web site, structural properties of the interface, channels and codes, types of menus and links. Functionality, considered the driving force of the evolution, was assessed according to the levels proposed by Shepherd and Watters (1998).

The evaluation has shown that the encyclopedias available online represent a collection of multifarious interfaces and forms offering different content. Similarities among encyclopedias involving the features of function, content, form and functionality enable the isolation of four main types of encyclopedia, representing different degrees of modification affecting the traditional generic properties and various stages of evolution and adaptation to the conditions of the medium. Therefore, we propose the following taxonomy of online encyclopedias, encompassing the main types and sub-types of encyclopedic sites:



Naturally, the proposed taxonomy does not constitute the one and only classification of online encyclopedias, as the medium creates favourable conditions for the formation of borderline constructions which defy clear-cut categorization. Yet, this taxonomy reflects the noticeable and popular tendencies concerning online encyclopedias, presenting the most conspicuous characteristics of the respective types.

What follows is the enumeration of the basic conventions concerning the encyclopedias:

4.2.1. Reproduced encyclopedias

A substantial percentage of web sites from the analysed collection (28 web sites – 35.90%) fell within the group of reproduced encyclopedias, which constitute examples of a transfer of the properties of a print encyclopedia to the World Wide Web, without substantial modifications of the genre.

Function

Reproduced encyclopedias, being direct replicas of paper antecedents, have predominantly an informative function.

Content

The content of reproduced online encyclopedias consists mainly of short articles presented alphabetically, in a linear order. Out of the whole corpus, only 4 encyclopedias offered thematic categorization.

Form

The form of a reproduced encyclopedia, apart from the textual body includes a menu providing links to other pages and a limited number of photographs. Nearly 60% of the reproduced encyclopedias from the analysed group (16 sites) were purely textual in form, the remaining additionally providing images or graphic items. The discourse properties of reproduced encyclopedias are equivalent to printed works.

Functionality

The functionality of such encyclopedias is limited, being restricted mainly to hyper-text and simple browsing options. See Fig.1: p. 40.

One search option

Article reproduction, no hyperlinks

The screenshot shows a web browser window with the address bar displaying 'http://www.bartleby.com/66/cy/Cyzicus.html'. The page features a navigation bar with links for 'Home', 'Subjects', 'Titles', and 'Authors'. Below this is a search bar and a list of reference works including 'Encyclopedia', 'Bibliography', 'Thesaurus', 'Quotations', and 'English Usage'. The main content area is titled 'Cyzicus' and includes a paragraph of text about the city's history. The text is reproduced verbatim from the source, with no hyperlinks. The page also includes a sidebar with 'Harry Potter' and 'Amazon.com' links, and a footer with 'Zalozhczko'.

Harry Potter
amazon.com

Bartleby.com
Great Books Online
Search:

Home | Subjects | Titles | Authors

Reference > Columbia Encyclopedia

< PREVIOUS

CONTENTS - INDEX - GUIDE - BIBLIOGRAPHIC RECORD

The Columbia Encyclopedia, Sixth Edition. 2001-07.

Cyzicus

(SIE-TEE) (KEO) - ancient city, NW Turkey, at the neck of the Cyzicus Peninsula. Strategically located, it ranked Byzantium in commercial importance. Founded (756 B.C.) by Greek colonists from Miletus, it was a member of the Delian League. In 410 B.C. Alcibiades defeated a Spartan fleet off Cyzicus, and in 74 B.C. the city withstood a siege by Mithradates VI of Pontus. As a reward for its loyalty, it became a free city under Roman rule. In A.D. 675 the town was pillaged by the Arabs and then used as a base for a siege of Constantinople. There are ruins of a large temple built by Roman Emperor Hadrian and of other public buildings.

The Columbia Encyclopedia, Sixth Edition. Copyright © 2001 Columbia University Press.

CONTENTS - INDEX - GUIDE - BIBLIOGRAPHIC RECORD

Zalozhczko

1. Fig. An instance of a reproduced encyclopedia (www.bartleby.com, 14th March 2008)

4.2.2. Adapted encyclopedias

Adapted encyclopedias, as the name suggests, have undergone a process of adaptation to the new medium in which they occur, utilizing the possibilities it offers. Still, the degree of enhancement and the scope of adaptation among the sites are noticeably uneven which necessitated the isolation of two sub-types of adapted encyclopedias, i.e. lower and advanced encyclopedias. The most important differentiating factor between them is the scope of the functionality attribute.

4.2.2.1. Lower adapted encyclopedias

This group of encyclopedias comprises 21.80% pages (17 web sites) which, in comparison to the previous collection of reproduced encyclopedias, offer additional content and an extended form.

Function

The informative purpose of adapted encyclopedias is enriched by a further function, encompassing interactivity enhancement.

Content

Apart from the textual body of articles with bibliographic records, lower adapted encyclopedias supply a range of indexing types, not only alphabetical, but also thematic, by date or by person. Moreover, below each article the user can find a list of articles thematically connected with the entry in question. Besides, lower adapted encyclopedias offer additional services, which include lists of featured articles, most popular entries or “articles of the day”. What is most important, these encyclopedias are continually updated and revised.

Form

The form of lower adapted encyclopedias, apart from the textual body, offers collections of photographs, and five of the sites included in this group give the option of using simple audio files. Such encyclopedias provide as well an extended menu including internal links to the content within the encyclopedia. As far as discourse properties are concerned, similarly to reproduced encyclopedias, lower adapted encyclopedias exhibit features analogous to the universal print encyclopedias.

Functionality

In addition to the functionality offered by the browsers, such as hyperlinking and searching, these encyclopedias offer a random choice facility, which gives the reader an article chosen at random by the browser. Besides, such encyclopedias increase the level of interactivity with the reader by offering short interactive quizzes in which users can test their knowledge on a particular area of study. This usually includes one question per day, concerning different disciplines of study. See Fig. 2: p. 46.

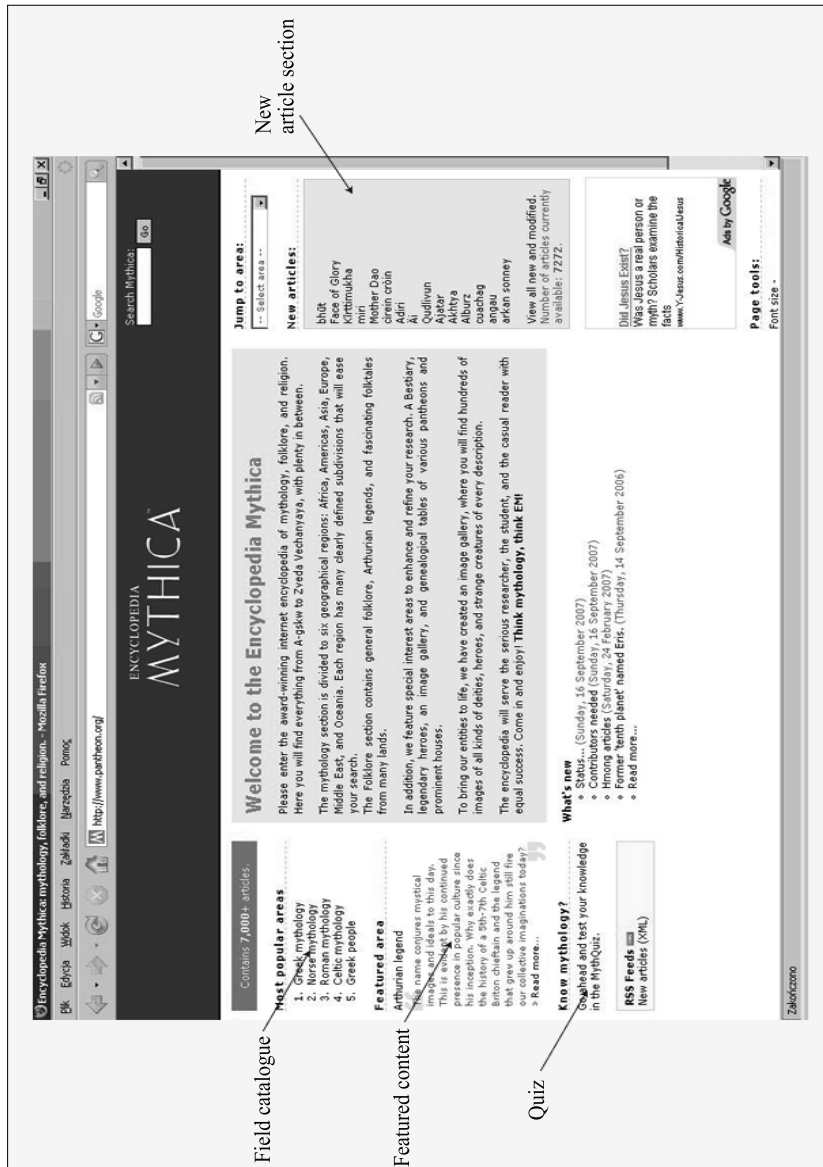


Fig. 2. An example of a lower adapted encyclopedia (www.pantheon.org, 14th March 2008)

4.2.2.2 Advanced encyclopedias

The collection of advanced encyclopedias, comprising 11.53% (nine web sites) of the whole collection, exhibits a far more extended degree of functionality and numerous additional services. The interface of such encyclopedias differs considerably from reproduced and lower adapted types. Such encyclopedias turn to the reader not only with the written word, but also with the photograph or other visual images, reproductions plus animations and sound. We might assume that advanced encyclopedias are progressing towards the interface of a multi-functional and multi-media portal, which can be observed in the content they offer.

Function

As far as advanced encyclopedias are concerned, the entertainment and interactivity-enhancing functions are superior, with the sites offering the possibility to contribute, though not without restrictions, as is the case with emergent encyclopedias.

Content

Advanced encyclopedias do not concentrate only on the presentation of encyclopedic articles, as can be observed in the substantial increase in the amount of content offered. Such pages, in addition to the content of the two previously mentioned types of encyclopedias, offer a number of additional services, such as a news stream and the presentation of important dates (timeline services, calendars). These encyclopedias offer a high level of customization. For instance, the Encyclopedia of Earth online (www.eoearth.org) offers articles differing in their scope of data coverage, from brief and simplified versions to complex treatment of the subject. What is more, users may find links to other projects, such as the blogs of editors or newsletters as well as other portals touching upon related materials or offering in-depth discussions of particular subjects.

Form

The form of advanced encyclopedias, apart from textual content and photographs, makes use of high-quality animations and video files, banners and a rich menu providing internal links to highly advanced galleries as well as external links leading to further sources (papers, presentations, archives, guides). The Encyclopedia of Life (www.eol.org), the most advanced encyclopedia within the analysed material, offers high-quality documentaries concerning life within the plant and animal world. Stylistically, these encyclopedias are alike the standard universal encyclopedias.

Functionality

The functionality offered by such encyclopedias is highly advanced. It does not only focus on browsing or hyperlinking, but offers a number of interactive services. Thus, users may not only browse through materials, but also download them, preview an animation, personalize the content (get shorter and simpler versions of articles, have materials from relevant fields sent via e-mail). What is more, the interactivity of such encyclopedias includes not only quizzes but, more importantly, also establishes contact with the board of editors of the encyclopedia via e-mail or a discussion forum. The users of the encyclopedia are invited to send their comments on particular articles and propose changes. Moreover, such encyclopedias allow a live-chat with experts in a particular field (cf. Encyclopedia of Life). See Fig. 3: p. 48, Fig. 4: p. 49, Fig. 5: p. 50.

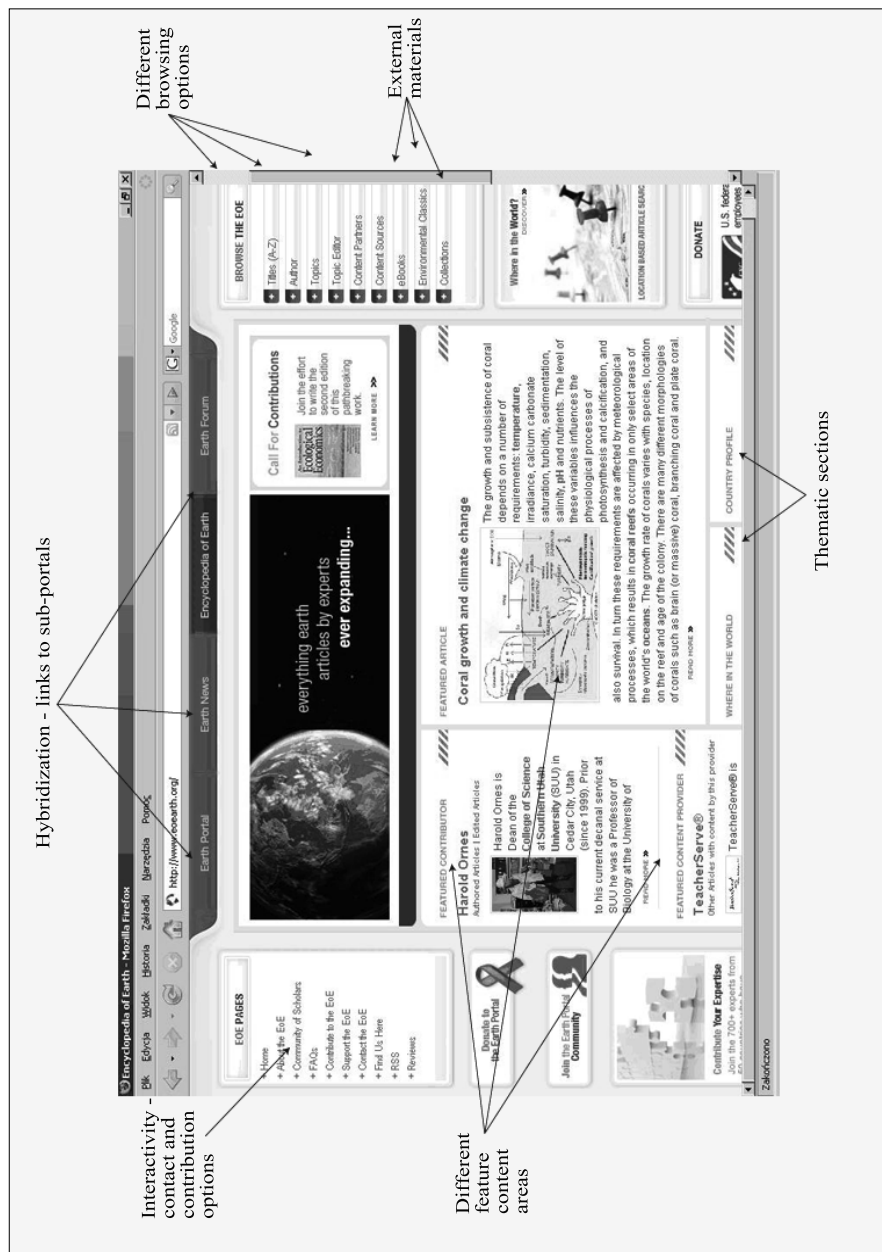


Fig. 3. An instance of an advanced encyclopedia (www.eoearth.org , 14th March 2008)

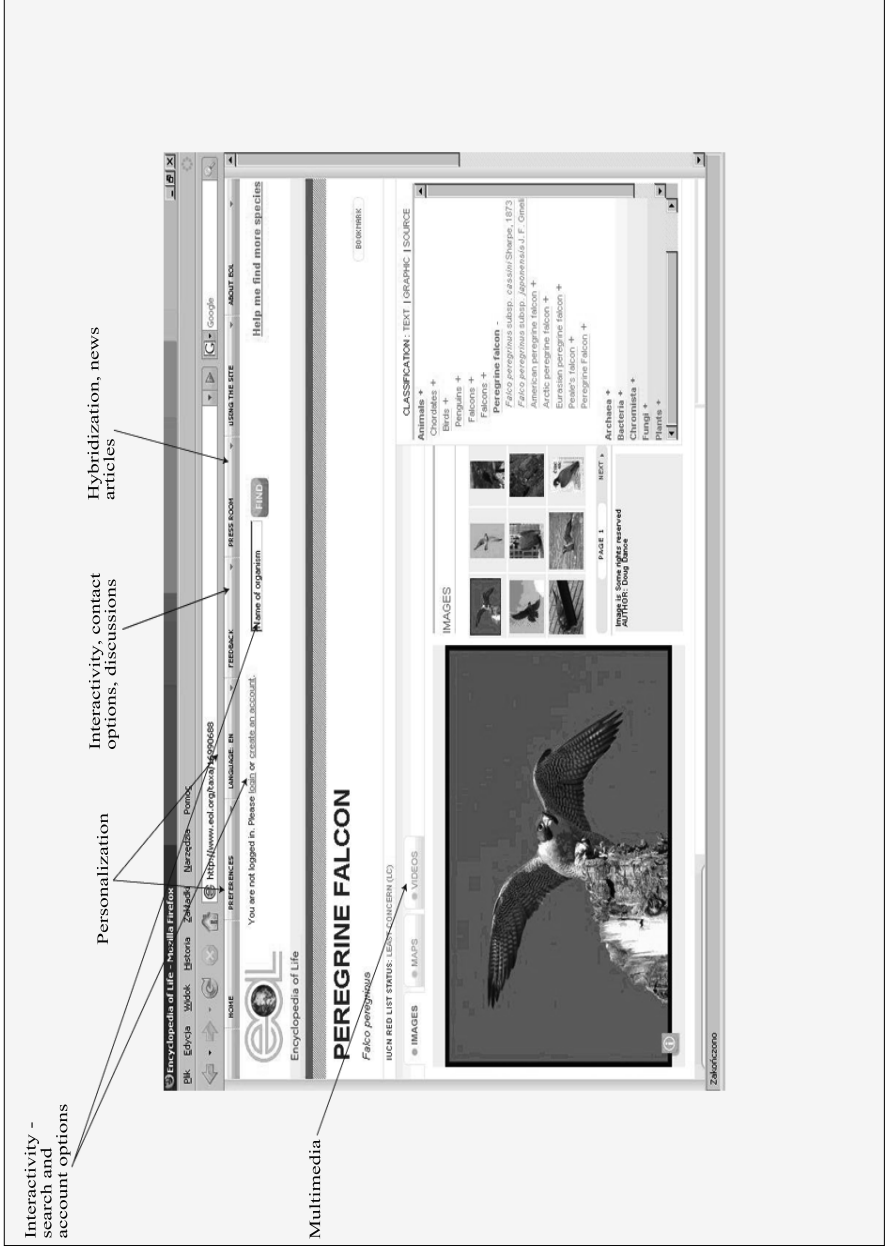


Fig. 4. An instance of an advanced encyclopedia (www.eol.org, 14th March 2008)

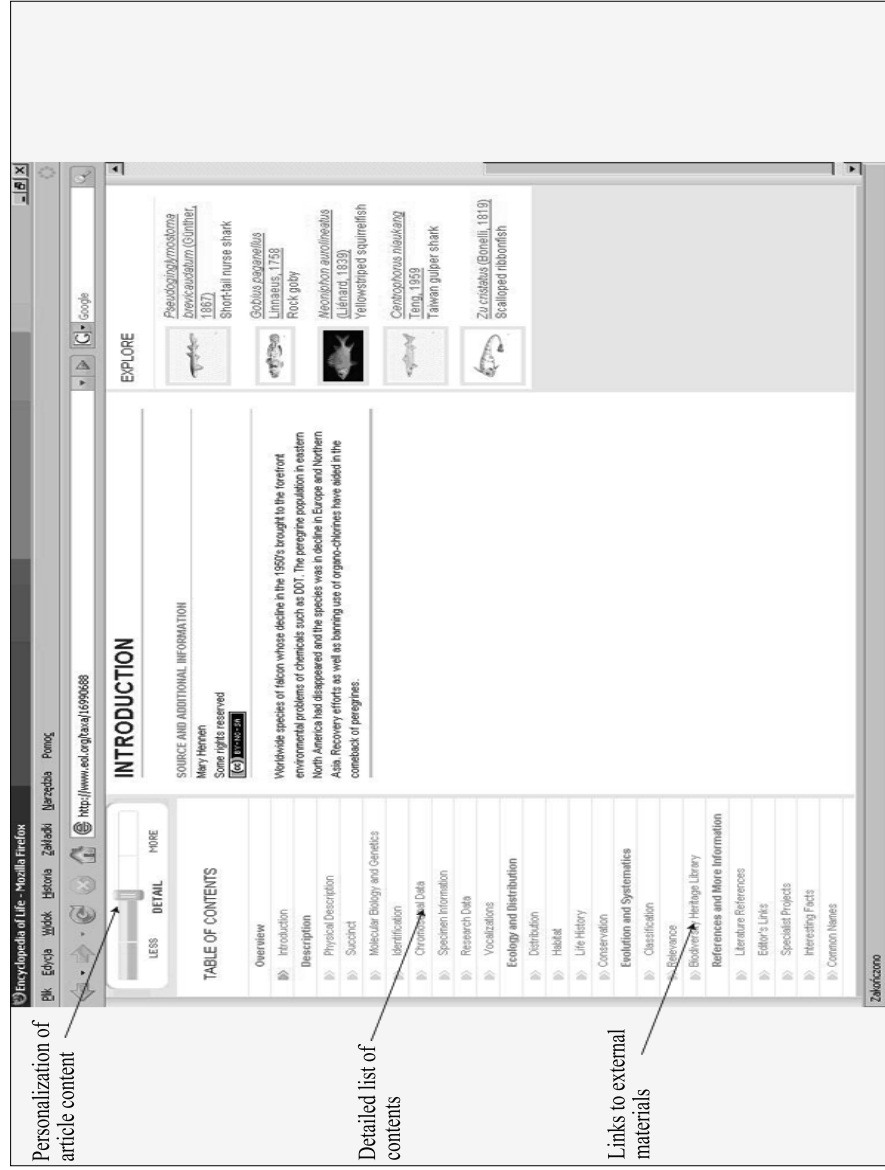


Fig. 5. Article structure in an advanced encyclopedia (www.eol.org, 14th March 2008)

4.3. Emergent encyclopedias

In this group of encyclopedias, comprising 23.07% (18 sites) of the whole collection, the evolution is based on a high degree of generic diversification entailing extensive transformations in structural, pragmatic as well as discourse aspects. This group comprises two main types of portals:

User-created portals – (14.10%) collaboratively created encyclopedic web pages, with diversified content extending the scope of an encyclopedic resource.

Resource pages – (8.97%) thematically oriented portals, presenting various types of materials from a specific domain of knowledge.

4.3.1. User-created portals

New information technologies underlying the birth of Web 2.0 have initiated the development of collaborative user-created projects concentrating on the production and dissemination of knowledge. The uniqueness of such projects lies in their openness to all internet users and a special emphasis put on egalitarianism and user-interaction.

Collaborative reference projects constitute an appealing subject for research, requiring a separate in-depth discussion, which is not possible here due to reasons of space. Below, we enumerate only the basic characteristics of such sites.

Function

As far as emergent user-created encyclopedias are concerned, the informative function is overwhelmed by a set of other roles the sites fulfil. The sites offer a high level of interactivity, which encompasses the possibility of user-contribution, encouraging the expression of personal views and exchange of opinions. These sites have an entertainment function as well, offering not only information, but also multi-media and interactive materials.

Content

The content of emergent user-created encyclopedias is not limited to encyclopedic articles, but encompasses various types of materials. As the introduction to one of these sites claims, it is “a growing resource that describes, tracks and explains culture, commerce, politics, media, sports, brands – everything possible, really” (<http://www.thedesignencyclopedia.org>, 18th March 2008).

Similarly, the invitation to an encyclopedia on the BBC portal states:

Don't panic - the most reassuring words you're likely to read all day. There really is no reason to panic, because h2g2 is your guide to Life, the Universe and, well, Everything. We have thousands of entries on all manner of subjects, all written by people like you. If we don't have the one you want, you can even write it yourself (<http://www.bbc.co.uk/dna/h2g2>, 14th March 2008).

The content includes apart from encyclopedic material also different ways of classification and indexing as well as links, multi-media, user comments, user dialogues and additional pages (galleries, supplements).

Form

Emergent user-created encyclopedic sites have the form of a collection of user-created articles associated with wide-ranging issues, a fact which is manifested in the introduction to one of the sites: “Everything2 is a collection of user-submitted writings about, well, pretty much everything” (www.everything2.com, 28th March 2008).

In many cases (7 sites), these projects are organized around a post paradigm, in contrast to other types built around a page paradigm. The post paradigm resembles small chunks of information, acting as input in a larger framework on the site.

Articles in these encyclopedias are accompanied by comments provided by other users, each in a separate box, with attribution details (nickname of a user and time of posting). These comments take the form of lists of short texts, usually two or three sentences long, ordered chronologically from the latest to the earliest.

Discourse properties of such sites deserve further attention and an in-depth discussion, since they constitute an interesting hybrid form, being a mixture of conventional encyclopedic articles, with colloquial speech-like posts.

Functionality

The most significant property of user-created encyclopedias is the opportunity for an individual unrestricted contribution to the encyclopedia that any user can make. This property constitutes the main reason for the departure of the genre from conventional forms, the result of individualization of the content based on user creativity, needs and wishes. Additionally, functionality comprises interactivity (e-mail options, chats), downloading, multi-media and content personalization. See Fig. 6: p. 53, Fig. 7: p. 54, Fig. 8: p. 55, Fig. 9: p. 56.



Fig. 6. An example of an emergent user-created encyclopedia (www.bbc.co.uk/dna/h2g2, 19th March 2008)

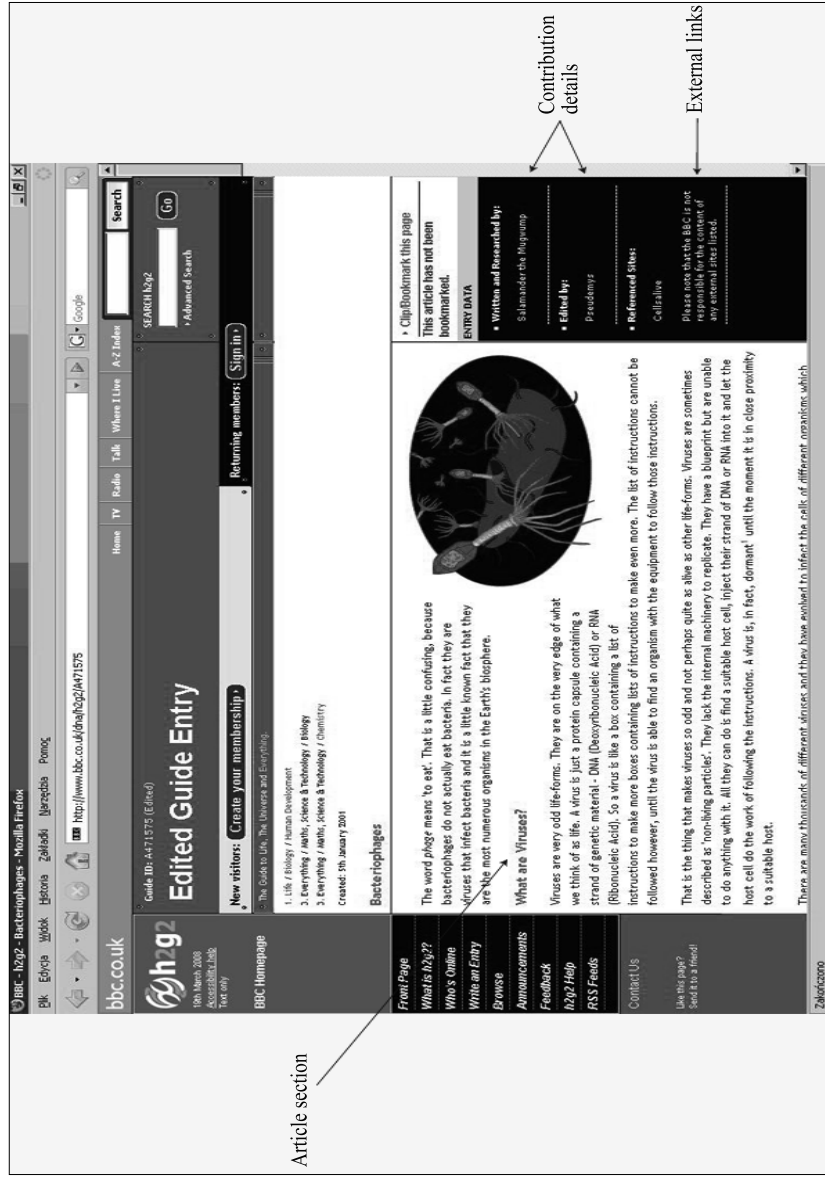


Fig. 8. Article structure in an emergent user-created encyclopedia (www.bbc.co.uk/dna/h2g2, 19th March 2008)

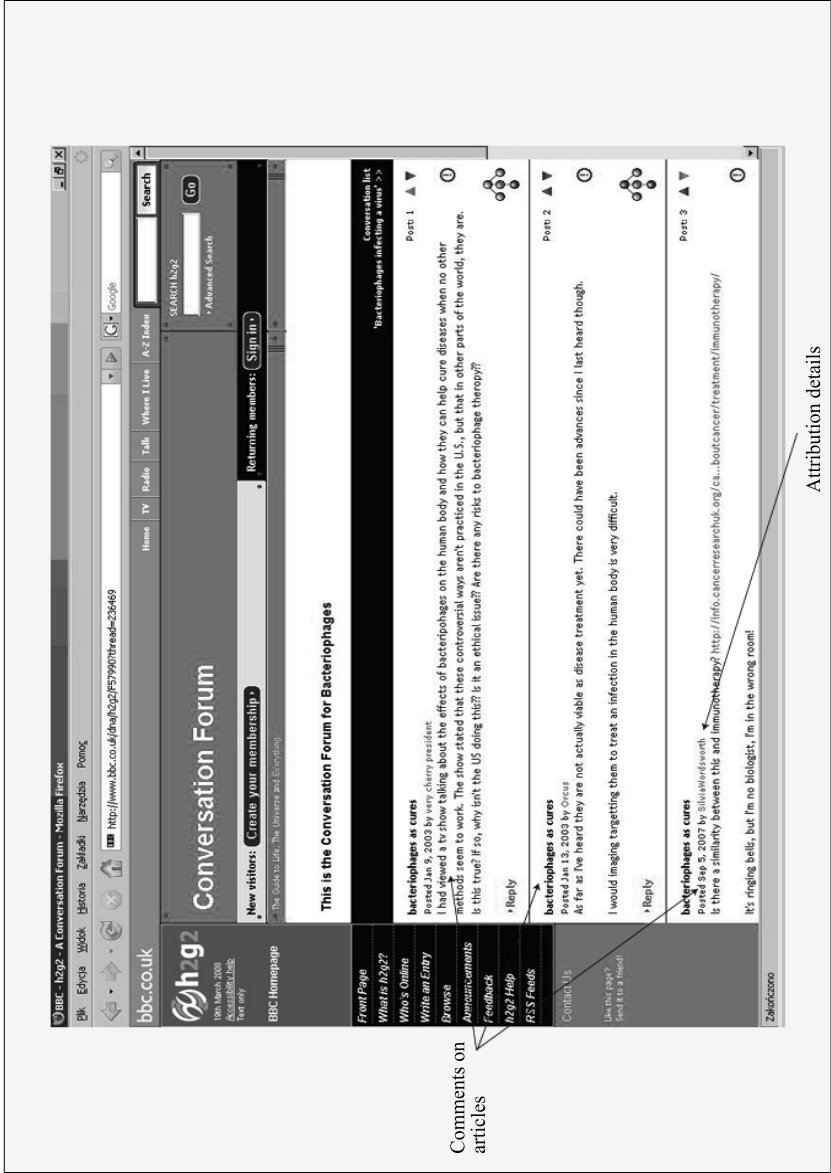


Fig. 9. User-comments on an article in an emergent user-created encyclopedia (www.bbc.co.uk/dna/h2g2, 19th March 2008)

4.3.2. Resource pages

Function

Resource pages comprise informative, educational and interaction-enhancing sites, where frequently knowledge acquisition is found in tandem with play and entertainment. These sites are predominantly children- or pupils-oriented, functioning as learning bases and explain issues associated with the school-curriculum.

Content

Resource pages comprise either single or multi-topic content. Single topic pages concentrate on one subject area, such as mathematics, history or music. The content of resource pages includes explanatory materials in the form of descriptive entries and in a question-answer format as well as games, puns and quizzes.

Form

Resource pages represent a variety of forms. Textual material is rather scarce, as the pages are dominated by graphic and animated files. Resource pages offer catalogues and a menu to the content as well as external links to further associated pages.

Functionality

Resource pages utilize the functionality of the medium in that they bring into one site a range of material, as in advanced multi-media encyclopedias, yet not only concentrate on informative entries, but involve also play and games. See Fig. 10: p. 58.

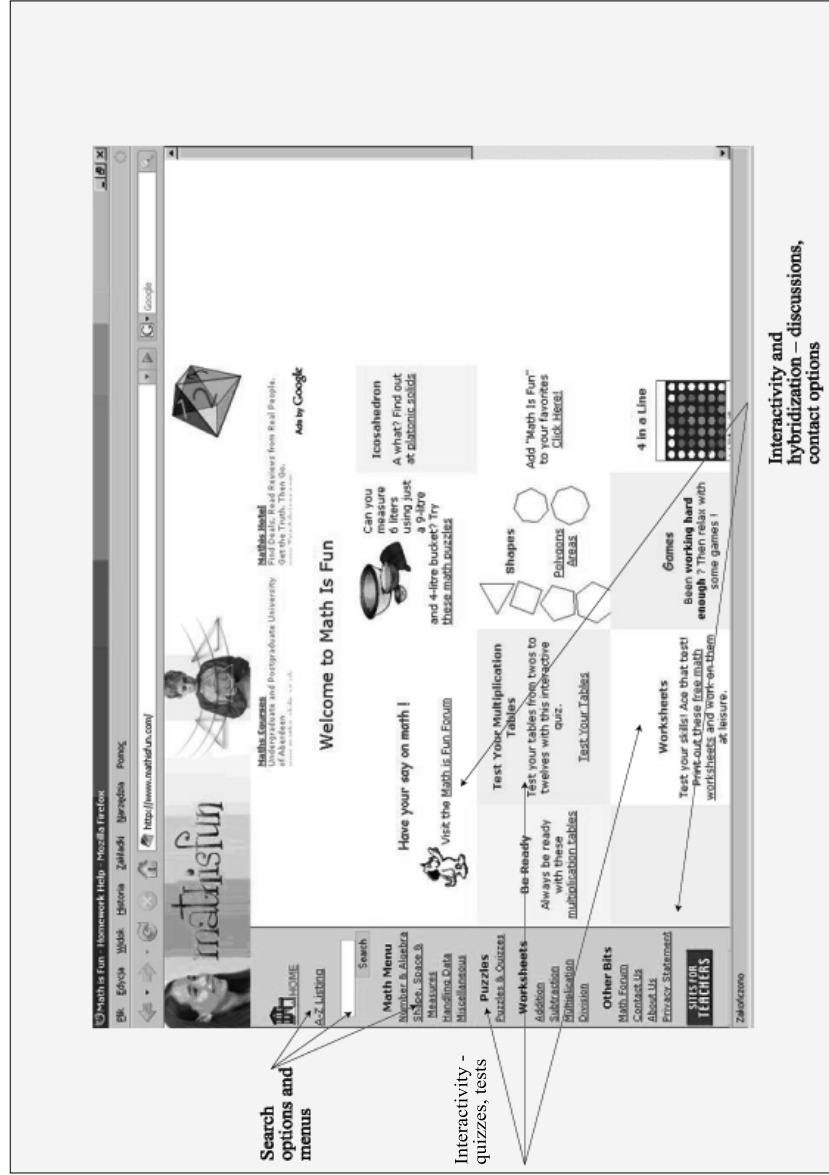


Fig. 10. An example of a thematic resource page (www.mathisfun.com, 14th March 2008)

4.4. Emerging encyclopedias

The encyclopedias investigated allow the isolation of a further type of reference sites, i.e. emerging encyclopedias, which depart considerably from the conventions associated with the genre of encyclopedia, as their properties evolve entirely from the affordances of the medium. This group comprises six sites, constituting 7.70% of the whole collection.

Function

The purpose of emerging sites is predominantly to enhance interactivity among internet users. What is more, these sites manifest a propensity towards an expressive function, being in the majority products of individual authors, thus reflecting their own interests and hobbies.

Content

Half of these web sites were created by individual authors, with the aim of presenting content of particular interest to themselves. This can be seen in the invitation presented by the author of the Freesoft page:

Welcome! The internet Encyclopedia is my attempt to take the internet tradition of open, free protocol specifications, merge it with a 1990s Web presentation, and produce a readable reference to the technical operation of the internet (<http://www.freesoft.org/CIE/>, 20th March 2008).

The content of such encyclopedias does not allow easy description, since they contain a diversified collection of materials, from web site addresses, to internet courses, and downloadable software.

Form

Emerging encyclopedias do not offer encyclopedic articles but comprise lists of various web site addresses referring to a variety of content and touching upon a range of subjects. For instance, as the headline of the Digital Universe encyclopedia states: “it is an ever-growing array of commercial-free portals mapping the highest-quality internet destinations” (<http://www.digitaluniverse.net>, 22nd March 2008).

What is interesting, the links representing a wide range of subjects may be perceived as an author’s internet surfing history record. The links offered by these sites may be followed by an author’s commentary concerning the content of the linked page.

Functionality

The functionality of emerging encyclopedias encompasses downloading, interactivity, contribution and use of multi-media.

This sub-type of online encyclopedias is a clear example of an individual approach to the genre, its content and form, constituting a hybrid mixture of different materials. Such constructions do not exhibit typical properties of the genre, signifying a considerable departure from the convention. See Fig. 11: p. 60, Fig. 12: p. 61.

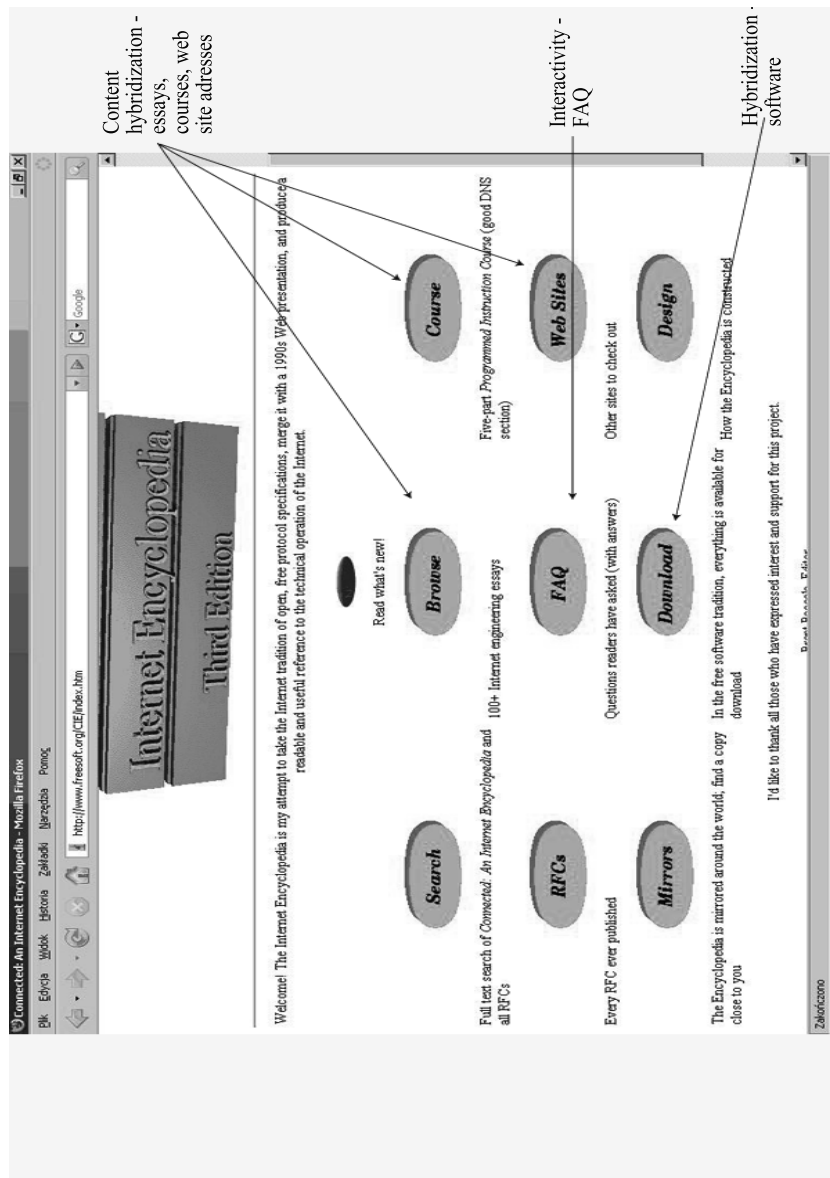


Fig. 11. An instance of an emerging encyclopedia – an example of hybridization and individualization of the site (www.freessoft.org, 20th March 2008)

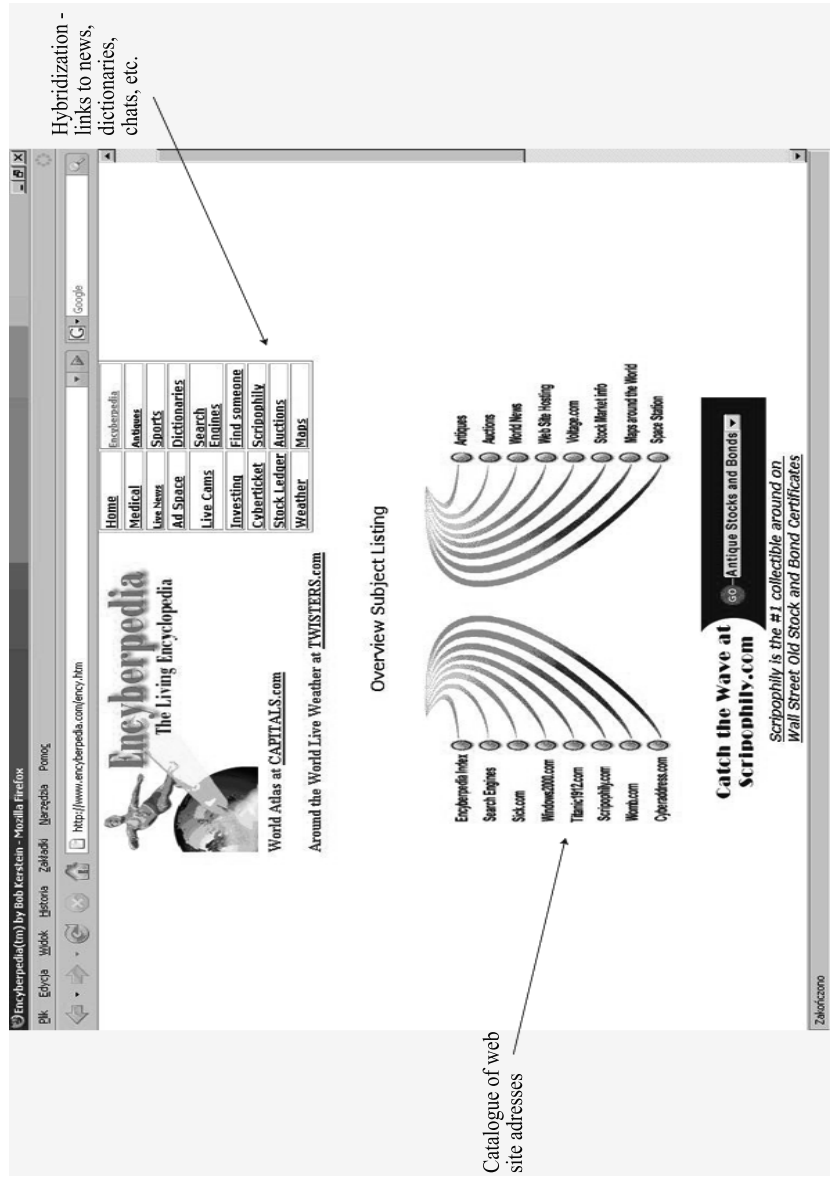


Fig. 12. A further instance of an emerging encyclopedia (www.encyberpedia.com, 20th March 2008)

The development of each of the types of online encyclopedia reflects the interplay of different processes accompanying the transfer of the genre to a new environment:

- Reproduced encyclopedias exemplify the process of replication of the genre on the web.
- Adapted encyclopedias, in addition to genre replication, undergo a process of enhancement. The difference between the two sub-types of adapted encyclopedias concerns the degree of enhancement, from relatively little with respect to lower adapted encyclopedias, encompassing only a few additional properties obtained from the new context, to considerably more as far as advanced encyclopedias are concerned, entailing the enrichment of the properties of function, content, form and functionality.
- Emergent encyclopedias develop under the influence of the processes of enhancement, hybridization and modification concerning the constituent parts of the generic structure.
- Emerging encyclopedias reflect the interplay of the processes of hybridization, modification and, most importantly, individualization on the generic form.

Moreover, what deserves particular attention and additionally differentiates novel online encyclopedias from extant types are the new names given to such structures. While all the reproduced and adapted pages use the name “encyclopedia” in their titles (e.g. Encyclopedia of Astronomy, Encyclopedia of Physics, New Georgia Encyclopedia), novel genres may be distinguished by innovative and original titles. Thus, among emergent sites we find a Toonopedia, a Scholarpedia or an Uncyclopedia, which combine the suffix *-pedia* from *encyclopedia*, with a different word, associated with either the content of the site (Toonopedia – about music, Uncyclopedia – containing absurd information) or the manner of composition (Scholarpedia – composed by experts). Among the emerging genres we find novel labels such as Encyberpedia or Webopedia clearly pointing to the association of the site with the medium of the internet.

A comparison of the main differences between particular types of online encyclopedias is presented in Table 1.

Properties	Replicated 35.90%	Adapted		Emergent			Emerging 7.70%
		Lower 21.80%	Advanced 11.53%	User-created 14.10%	Resource pages 8.97%		
Function	informative	informative, interactive	informative, interactive, entertainment	interactive, informative, entertainment, expressive	informative, educational, interactive, entertainment	interactive, entertainment, expressive, informative	
Content	short texts presenting encyclopedic information	encyclopedic entries, featured articles, articles of the day, quizzes, calendar	Encyclopedic articles, featured content, news, material prepared especially for the web, updates, scholarly papers, informational and educational texts, blogs, newsletters, commentaries	user-written posts of various length covering different topics, encyclopedic materials, newsletters, discussions, user-personal spaces	textual and graphic materials, animations, puzzles, games, lesson-tasks	list of links to other portals with or without description	
Form	textual body, limited amount of photographs	alphabetic and thematic categorization, photographs, simple menus	considerable growth of the length of the web site, graphic, audio and video files, extended menus, external and internal links	list of user-posts, menus, internal and external links, textual as well as graphic and video files, different catalogues	collection of menus to different materials, e.g. definitions, tests, descriptions, external and internal links	form of a list or catalogue of web site addresses	
Functionality	browsing and searching	e-mail contact, random choice, interactive quizzes	interactivity, galleries, discussion, login function, online chat, personalization, portals, animations, downloading	interactivity, contribution, login function, e-mail options, downloading, personalization	browsing, searching, interactivity, downloading, login function	browsing, interactivity, downloading	

Table 1. The comparison of online encyclopedias

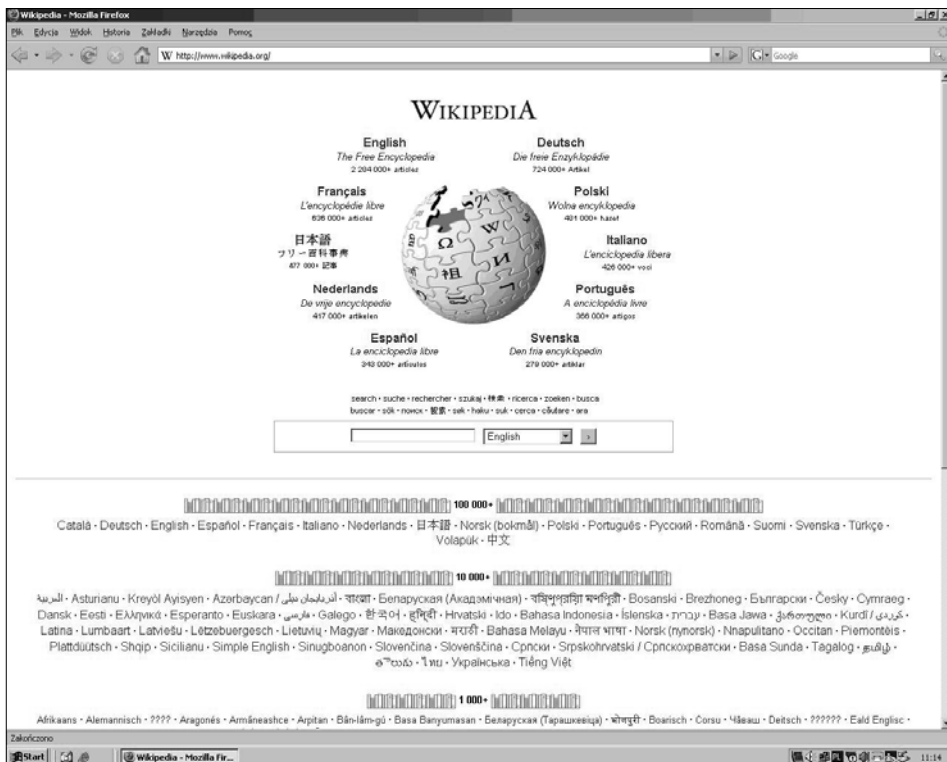
To conclude, the encyclopedias analysed prove that online works of reference represent a wide diversity of generic structures, which exploit the new affordances of the medium to various extents, thus leading to the emergence of innovative encyclopedic forms.

The comparison between online encyclopedias and their traditional equivalents lets us see that each type has their own advantages and disadvantages. Among the most significant advantages of online resources belong the convenience of use, accessibility, low cost, flexibility, interactivity, increase in content and storage capacity (Frost 2005). Yet, there are a few areas in which traditional paper antecedents still outweigh their modern equivalents. As with paper dictionaries, encyclopedias have a symbolic value, constituting important achievements closely associated with a particular culture (e.g. *Encyclopedia Britannica*, *L'Encyclopédie*, *ESPASA*). Besides, a further important aspect in favour of traditional reference is their durability. Due to the constant advancement of technology, electronic carriers become obsolete very quickly, while printed works may last for centuries (cf. Landau 2001, Frost 2005).

A summary of the main transformations triggered by the arrival of encyclopedias on the web involve changes in the following areas:

Cost	→	elimination of the costs of paper, printing and binding but involve costs of site maintenance
Price	→	free or by subscription
Speed of publication	→	technically faster
Interactivity	→	increased, easy feedback
Additional features	→	able to include recent data, animation, unlimited length of articles.

5. Wikipedia as a genre



65

Fig. 1. Wikipedia homepage – the gate to different language versions of the encyclopedia (http://www.wikipedia.org, 16th October 2007)

5.1. Wikipedia – introduction to the project

The analysis of the evolution of genres, as Witosz (2005) emphasizes, should take into account not only the presentation of the newly formed structures, or the discussion of disappearing genres and their relations to new forms, but it should also discover traces of the past or traces of continuation of older genres in those newly emerging. Since Wikipedia evolved from a print antecedent, it is relevant to study the interdependence between its structure and that of a print equivalent in order to observe whether, and if so, to what extent the medium and the technological developments influence the transformation of the genre. In other words, it is important to note to what degree Wikipedia reproduces its print antecedents, and what modifications are introduced into the traditional textual form of an encyclopedia.

Wikipedia constitutes the most appropriate subject for an analysis of the genre of online encyclopedias for a number of reasons. Firstly, while there are a number of other encyclopedias available online, many of which are online versions of print encyclopedias already well established in the market (e.g. *Encyclopedia Britannica*, *MSN Encarta* or *Columbia Encyclopedia*), Wikipedia is the most popular, the most frequently visited and referred to online encyclopedia.¹ This encyclopedia is recognized internationally, having developed equivalents in more than 250 language versions, in contrast to other online encyclopedias. Secondly, Wikipedia constitutes a product of the web, of internet technology, and exploits all the possibilities afforded by the medium, which are believed to exert a profound influence on the structure of a genre (Crowston and Williams 1999, Witosz 2005). Third, Wikipedia constitutes the most extraordinary phenomenon on the web, having gathered millions of volunteers in the compilation of encyclopedic articles. Unlike other online encyclopedias it allows everyone to create articles freely, without restrictions.

The following chapter first describes the structural features of Wikipedia, its construction and mechanisms of operation. Next, Wikipedia is compared with a print encyclopedia in order to determine the genre characteristics of Wikipedia and trace the evolution of the genre from print to the web.

5.1.1. Wikipedia mechanism

Wikipedia – the free encyclopedia was founded in 2001 by Larry Sanger and Jimmy Wales as an alternative to Nupedia, a strictly controlled online encyclopedia written by experts (basically by PhD students). The project of Nupedia was not successful and failed with only a few hundred entries, while Wikipedia developed quickly and within a short period of time became the largest reference site on the web. The whole project comprises (on February 2008) nearly 9,558,178 articles, written in 256 languages.²

¹ Since January 2006, average daily page view score of Britannica's site constituted 1% of Wikipedia's. Alexa.com (a web site computing the popularity of web sites) observed more than 2000 accesses to Wikipedia a second, which places the encyclopedia among the 20 most popular internet pages (www.en.wikipedia.org)

² Data from www.en.wikipedia.org

Wikipedia is based on a unique system – a wiki-mechanism, created in 1995 by Ward Cunningham. This mechanism enables any user to modify web page contents or to create new pages freely, using any of the available internet browsers. Due to this system, in Wikipedia any user can introduce new entries, edit and modify the articles already present in the encyclopedia. The name *wiki* is of Hawaiian origin, where it means “quickly”, referring to the speed with which users can edit or modify texts on the web.

The wiki-mechanism is considered a significant innovation in that it has enabled collaborative authoring and greatly facilitated the edition of web pages (Desilets *et al.* 2005: 1). This system distinguishes wiki sites from other typical web sites, where the network cannot be modified by external persons, as it remains under supervision of its owners (Capocci *et al.* 2006: 1). In a wiki-system, each page has an “Edit this page” link which displays the text ready for modification in an editable field. To edit a page users have to follow a simple procedure of clicking on this node, modifying the text and then submitting the changes using the “Save” button. The system itself converts the text into an HTML format. Such editing of pages is possible in almost all parts of the site. Moreover, every single change is automatically saved by the system. Wiki also facilitates the creation of links between pages. To create links, users need to open the page where the link is to be inserted and type the name of the page that is to be indicated, using a specific syntax called WikiWords (e.g. SomePage, some_page). The procedure is finalized upon clicking on the “Save” button, analogous to creating new content.

Its effectiveness and the relative ease of its use constitute the main reasons behind the popularity of the wiki-mechanism. The evidence of this usability is visible in the huge number of authors editing Wikipedia. A further proof was provided by the research of Desilets *et al.* (2005), who investigated the usage of the system among children and indicated that the mechanism can be successfully managed by technically inept users.

The wiki-mechanism has gained widespread popularity among internet users.³ It serves as a basis for other collaborative projects, devoted to various topics, ranging from literature, through film and music to virtual games.⁴

5.1.2. Dimensions of the project

It can be argued that Wikipedia has become a global and largely international product, including the editions in Lombard, Cebuano, Volapuk and Yoruba.⁵ The existence of so many versions of Wikipedia constitutes an indubitable and exceptional proof of its popularity and reflects the success of the project.

³ More information on wiki-sites can be found at: http://personalweb.about.com/od/wikihostingandsoftware/Wiki_Sites_Wikia_Wikicities_etc.htm

⁴ A list of different wiki-sites is available at: http://www.wikia.com/wiki/List_of_Wikia

⁵ The data come from Wikipedia Statistics (February 2008) available at: http://en.wikipedia.org/wiki/Wikipedia:Multilingual_statistics

Wikipedias in different languages function on an autonomous basis and differ in size as well as degree of development. The largest Wikipedia is in English, counting over 2, 200 000 articles. Two others, German and French, exceed 600,000 articles, whereas Italian, Polish, Japanese, Dutch and Portuguese contain more than 400,000 entries. Every day around 1,600 articles are added to the English version of the encyclopedia, whereas the already existing entries are extended and modified on a continuous basis. The following Table presents the exact count of the largest Wikipedias, illustrating the number of articles and their size:

	All Wikipedias	English Wikipedia	German Wikipedia	Polish Wikipedia
Number of articles	9,558,178	2,203,380	700,983	465,678
New articles per month	255,529	47,690	14,111	11,073
Size of articles (in bytes)	2157	3141	2876	1375
Number of edits per article	20.4	33.5	28.7	10.9

Table 1. Number of articles in the largest Wikipedias

The findings of the research carried out by Zlatic *et al.* (2006: 1) prove that all Wikipedias, though autonomous, realize the same universal process of development despite being implemented by different communities. It has been observed (*ibid.*) that the larger and more popular Wikipedias function as examples and sources of inspiration for smaller projects. The interdependence between Wikipedias arises also from the fact that authors frequently co-operate in the production of different Wikipedias. Moreover, the content and data are frequently being transferred and translated into other languages. It has to be noted, though, that each Wikipedia also has a set of unique characteristics, such as a different design and organization of the home page, different structure of the network or different rules governing community membership (Zlatic *et al.* 2006: 2).⁶ Differences in content between Wikipedias concern mainly the scope of data with the focus more on nation-specific information, especially with reference to cultural, political or historical matters.

⁶ Zlatic *et al.* (2006) found that Polish and Italian Wikipedias differ significantly from the other networks. The differences involve the degree of distributions, assortativity, clustering and the triad significance profile. The dissimilarity proved to be the greatest in the case of the Polish Wikipedia due increased of the calendar pages. Such research proves that the process of universal growth of Wikipedias may be influenced by community-driven decisions.

Wikipedia constitutes a highly structured mechanism itself. It is divided into nearly 20 sections, each of which serves a special purpose.⁷ Viegas *et al.* (2007) investigated the evolution of Wikipedia over two years from 2003 to 2005. As the scholars (*ibid.*) observed, the complexity of Wikipedia's structure tends to increase dynamically. Their research proved that apart from the number of articles, the number of sites devoted to the management of the project has considerably increased.

The Wikipedia network, though complex, is characterized by a large interconnection of the links (Capocci *et al.* 2006: 2). Interconnection refers to the fact that by means of hyperlinks different pages, both internal and external to the project, can be easily accessed. The process of interlinking is facilitated by the simple procedure of the creation of new nodes, described above.

5.1.3. Legal issues

The content of Wikipedia is available to all internet users under the GNU⁸ Free Documentation License (GFDL). Free licenses, such as GNU, allow the use and distribution of computer applications without restrictions for any internet user. The introduction of the license was based on the conviction that "people should be free to use software in all the ways that are socially useful."⁹

Wikipedia is the largest collaborative environment using the GNU license. The license eliminates an individual approach to articles and strengthens the sense of collectivism (Holloway *et al.* 2005: 5). Wikipedians feel that the "authorship of data is irrelevant and sometimes even detrimental to the creation of truly communal repositories of knowledge."¹⁰ However, as Holloway *et al.* (*ibid.*) observe, the authorship of information does seem to be considered significant, which may be observed in authors' discussions on the talk pages.

The license also helped to popularize the encyclopedia, as any other site can mirror its text without special permission.

5.2. Wikipedia versus traditional encyclopedia

The analysis of its structure and the comparison of Wikipedia with a print antecedent of the genre allow us to distinguish a number of significant differences between the two instances of encyclopedias. The most profound modifications concern the follow-

⁷ E.g. Media (concerning audio and video files), Image (providing information about images), Category (referring to classification of contents), Portal (pages concerning specific topics), User (providing pages for registered users' personal presentation and auxiliary pages for personal use), Talk (directing to discussions concerning contents) and Help (providing information for new users on the workings of the wiki-mechanism)

⁸ GNU is a recursive acronym for "GNU's Not Unix" <http://www.gnu.org/>

⁹ See http://en.wikipedia.org/wiki/Main_Page

¹⁰ *Ibid.*

ing aspects: function, participant roles, content, form and functionality. It is in these areas that the influence of the medium is particularly extensive.

5.2.1. Function of Wikipedia

With the advent of new technologies, users' expectations towards reference works have changed. Users expect the information to be rich, attractive and dynamic. For such users, as Adams (2007: 9) observes, a hyperlinked text is more abundant and informative than the static traditional paper document and less constraining due to its non-linear structure – they can follow their interests and move through a wide network of texts upon clicking on the hyperlinks.

Wikipedia clearly exemplifies the extension and transformation of the functional aspect of the encyclopedia. Primarily, the free encyclopedia fulfils an informative role, as it constitutes a source of information covering a variety of subjects and fields of knowledge. However, in contrast to a traditional paper encyclopedia, Wikipedia holds as well practical-technical and didactic functions. It offers technical information describing not only the nature of a number of scientific processes, but also the usage and inner workings of various technologies.

What is the most interesting, though, is the expressive function that Wikipedia can also fulfil. The free encyclopedia gives its users the opportunity to express their views, exchange experiences and commentaries concerning not only particular entries in the encyclopedia, but also other subjects, not yet covered and explored. Wikipedia was the first to introduce the discussion entry which follows each article, where users, readers and authors can express their doubts or evaluations of the content. Due to its collaborative character, Wikipedia as well supports the creation and maintenance of contacts between contributors, encouraging discussions and cooperation.

Moreover, although Wikipedia does not pursue the ancient goal of the betterment of humankind, it does seem to embody more idealistic values, reaching beyond the dissemination of information. The ideals which Wikipedia promotes entail freedom of knowledge, cooperation between people, volunteerism as well as equal rights to knowledge production (Tereszkiewicz 2008).

5.2.2. Content

Wikipedia represents a contradiction of the traditional course of selecting and editing the encyclopedic content.

5.2.2.1. Authorship

The principles of authorship constitute the most significant factor differentiating the free encyclopedia from its paper antecedents. The opportunities for contribution offered by the site have changed the scope of traditional collaboration by awarding the role of authoritative editors to a group of volunteers. Such a change exerts a decisive impact on the structure and the way of functioning of the site.

The content of Wikipedia is created both by occasional contributors as well as regular authors. The whole Wikipedia project gathers as many as 222,000 different contributors, of whom nearly 115,000 individuals contribute to the English version of the free encyclopedia.¹¹ Contributors and users of Wikipedia can either remain anonymous or register in the project. Registered users obtain a user profile and are awarded more rights and greater contribution opportunities. The following Table presents the data regarding the number of contributors, active editors (editing more than 5 edits) and those who joined the project in May 2006:

	All Wikipedias	English Wikipedia	German Wikipedia	Polish Wikipedia
Contributors	222,187	114,893	32,768	5,248
Active contributors	63,974	35,836	6,837	1,380
New contributors	18,160	9,423	1,594	430

Table 2. Number of contributors in different Wikipedias

Wikipedia has undermined the traditional distribution of roles into those of a writer and a reader. The wiki-system gives the user the opportunity to become the author of the message, and thus every person can assume the roles of both writer and reader. The activity of the reader involves different levels, it progresses from reading the encyclopedia, through the possibility of changing or revising the messages, to the opportunity of composing new articles or deleting already existing entries. The identity of an author becomes diffused, as the text constitutes a product of a wide circle of users of the encyclopedia. Every reader consulting Wikipedia, upon finding the information inappropriate, irrelevant, inadequate or disagreeing with the conception of the issue can edit the entry and introduce any change s/he wishes, thus becoming an author of the entry. The shape of most Wikipedia articles constantly evolves. For instance, the entry on *United Nations* was composed over the period of six years by more than seven hundred authors who made more than four thousand edits.

The restatement of editorial conventions proposed by Wikipedia carries important implications for the idea of encyclopedia-making and our understanding of knowledge. Conventionally, an encyclopedic entry, produced by one person, depicted knowledge in a stable and one-dimensional manner. The reader was given only the end product of the editorial process and was not admitted into the course of entry-production. Wikipedia, on the other hand, shows that knowledge is debatable, that the way towards creating a final account requires argument and compromise.

¹¹ http://en.wikipedia.org/wiki/Wikipedia:Community_Portal

5.2.2.1.1. Wikipedia community

One of the mysteries of Wikipedia is the motivation of the volunteers who donate their time to the creation of the site. The authorship and management of the site are not rewarded in any material sense, as the contributors do not receive any remuneration for their work. Jimmy Wales, one of the founders of the project, characterizes the editors as “extremely information hungry, geeky kinds of persons.”¹² Adams (2007: 9) perceives them as “hyper-romantics who have conflated their imaginations and the scrapes of human knowledge into a hypertextual hyperuniverse.” The main incentive to contribute to the online encyclopedia seems to be a kind of satisfaction derived from the creation of a source of knowledge, a feeling of participation in a large and widely recognized project as well as of performing important editorial functions (Bryant *et al.* 2005). It has also been hypothesized that peer recognition plays an important role in attracting users to contribute to the encyclopedia (Kuznetsov 2006: 5).

The circle of contributors and authors forms an interesting internet community. The research on the community of Wikipedians (Bryant *et al.* 2005, Viegas *et al.* 2007) proves that with their growing involvement in the project, the roles of Wikipedians vary, as they undertake new activities serving different purposes.¹³

The so-called “Community Portal”, i.e. “the central place to find out what’s happening on Wikipedia.”¹⁴ This portal constitutes the “headquarters” of Wikipedia and includes the policies of the encyclopedia, editorial information, articles which need to be written or improved as well as collaboration opportunities, which unite different users around one entry in order to enhance its quality. Both Stivilia *et al.* (2005) and Viegas *et al.* (2007) emphasize its role in outlining guidelines for the authors’ conduct as well as editing activities. The community has its own bulletin board, where all news and announcements are presented.

The community of Wikipedia is highly structured. Users who gain a reputation in the community through their contributions constitute a “merit-based aristocracy” (Halloway *et al.* 2005: 6), with several layers of privilege: the higher level of administrators can grant the administrator status to others, the lower administrators have the right to delete and block pages, lower still are the regular users, and the lowest in rank, the anonymous contributors (*ibid.*). Proper collaboration between all the layers of the community is crucial as the content and structure of Wikipedia are complex and multifaceted. In practice, however, as the scholars note (Stivilia *et al.* 2005), the cooperation between the members of the community and the management of the encyclopedia are not always smooth and stratification leads very often to quarrels and editing wars between administrators.¹⁵

¹² Cited in Pullen (2005: 1).

¹³ Bryant *et al.* (2005: 4) outlined the path of contribution: novice users concentrate on correcting small mistakes or inaccuracies, then move toward fuller participation by taking care of a collection of articles, and, over time, with the growth of this collection, their goals expand, so they start to identify with the whole community and to perceive themselves as managers and creators.

¹⁴ http://en.wikipedia.org/wiki/Wikipedia:Community_Portal

¹⁵ Examples of such editing wars can be observed on the discussion pages accompanying each entry, or on the Wikipedia’s e-mail discussion lists.

5.2.2.2. Scope

Wikipedia, and nearly all online encyclopedias, denies the concept of exhaustiveness, owing to the lack of space restrictions.

Due to the editorial freedom and the constant development of the work it is practically impossible to determine the exact coverage of the encyclopedia. Wikipedia offers a largely assorted collection of articles and its thematic coverage is undoubtedly broader than that of traditional print encyclopedias. Wikipedia may be considered a composite work, combining the content of both a general, specialized and popular encyclopedia.

Specifically, Wikipedia comprises numerous entries describing in detail specific terms, especially from the field of computer science, physics and mathematics, exceeding the scope of a general reference work. Within the analysed material of 968 entries, we identified 194 extensive articles presenting a thorough discussion of specialized notions, e.g. articles on *Ubiquitin*, *Ubuntu*, *Uudecoding*.¹⁶

Wikipedia is particularly strong in the field of popular culture. Moreover, it offers entries which would not be considered encyclopedic in nature or content, such as descriptions of fictional characters, games, summaries of computer games or fiction, articles on famous sayings or quotes and on many uncommon phenomena. Within the analysed material, 245 articles investigated subjects associated with such topics, e.g. *Ubiklubi*, *Ufimtsev Defence*, *Ufo Aftershock*, *Umeshu*, *Under the Bridge*, *Uryu Ishida*, *Utumno*. Thus, we might say that Wikipedia forms a summary of all common and general encyclopedias, with a particular emphasis on popular themes.

Furthermore, what distinguishes Wikipedia from its traditional antecedents is a high number of how-to entries. The free encyclopedia provides instructions on the techniques of coffee preparation, underwater diving, pipe smoking, building a stove or curing hiccups (Schiff 2006).

The areas in which Wikipedia still falls behind the paper counterparts include mainly literature, philosophy and social sciences (Halavais and Lackaff 2008).

Wikipedia does not delimit the audience it is directed at, yet its coverage might indicate that it is targeted at a general, mainly adult audience. As Wallace and Fleet (2005) imply, owing to the collaborative character of the work, the primary audience may also encompass the contributors themselves.

5.2.2.3. Adjuncts

a) Multimedia

A further important feature which distinguishes Wikipedia from its print equivalents is the introduction of different codes, i.e. not only verbal, but also audiovisual. Wikipedia offers a large collection of images, illustrations, graphics and diagrams of various kinds as well as animations, all of which help to illustrate the phenomena presented in the textual form.

Although the inclusion of pictorial information has long been practiced in paper editions of encyclopedias, the fundamental difference between the free encyclopedia

¹⁶ As mentioned in the introduction, the analysis comprises articles beginning with letter U, chosen arbitrarily from the free encyclopedia.

and its antecedents concerns the number of such adjuncts and their variety. To exemplify this contrast we compared the number of illustrations in Wikipedia and in the 15th edition of the *Encyclopedia Britannica*. As contrasted with the print edition of the *Encyclopedia Britannica* (1990), Wikipedia includes three times more pictures and graphic files. Within 535 articles in Britannica's *Micropaedia* beginning with letter *U*, there were 78 pictures, 11 maps and 3 diagrams, while *Macropaedia* offers on average 7 pictures per article (e.g. in the entry on *Union of Soviet Republic* there are 15 maps, on *Uruguay* only 3). Within the analysed material of 968 articles in Wikipedia, 357 (37%) were not accompanied by any illustration, while 609 (63%) included images, 377 (62%) of which were photographs, 85 (14%) maps and diagrams and 146 (24%) scanned pictures or graphic representations.

As observed by Viegas (2007), the image namespace constitutes one of the fastest growing sections of Wikipedia. Such a tendency illustrates the growing importance of the visual side of the encyclopedia. Research shows that many users of Wikipedia concentrate solely on the contribution of images, finding it the most attractive aspect of collaboration (*ibid.*).¹⁷

The quality of visual content is, however, somewhat variable, ranging from scanned images to high-quality exact and professionally-taken pictorial representations. Visual materials, apart from enriching the content in an artistic and graphic respect, making them more attractive, fulfil primarily informative functions. Visual additions:

- constitute an element of description, where the text is insufficient (e.g. a photo or a picture of an animal or plant or work of art), e.g. Fig. 1: p. 75.
- complement the text, thus facilitating its comprehension due to a visual presentation of the content described, e.g. Fig. 2: p. 75.
- constitute collective combinations which facilitate the comprehension of the elements interspersed in the whole encyclopedia (e.g. Wikipedia's Timelines, political events);
- illustrations, esp. maps may constitute a document in themselves. Such illustrations can be read as a text, as they provide both data and a pictorial representation (e.g. political or economic maps) (Mielczarek 1972), e.g. Fig. 3: p.76.

As Viegas (2007) claims, the fact that many articles still lack visual representation may be explained by the inner mechanism of the wiki technology, which was predominantly designed for manipulating texts, not images or graphic representations. This is due to the fact that the contribution of images, graphics and video files proves more complex than the management of the textual content (*ibid.*).

Image files are different from the textual content in the following aspects (after Viegas *ibid.*: 2):

- Contrarily to textual files, which are the basic part of the wiki pages, images constitute external, "attached" files that can be added to the site.

¹⁷ According to Viegas (2007), the group of image contributors comprises mainly middle-aged, male users from Europe and North America, working in the IT field. The reasons for contribution include the following: supply of images to articles lacking visual representations; improving the quality of existing images; a wish to make a long-lasting contribution; a love of photography. These reasons signify that it is the enhancement of the encyclopedic contents that propels contributors to add visual materials.



Fig. 1. *Urania riphaeus*
(http://en.wikipedia.org/wiki/Urania_riphaeus, 15th February 2007)

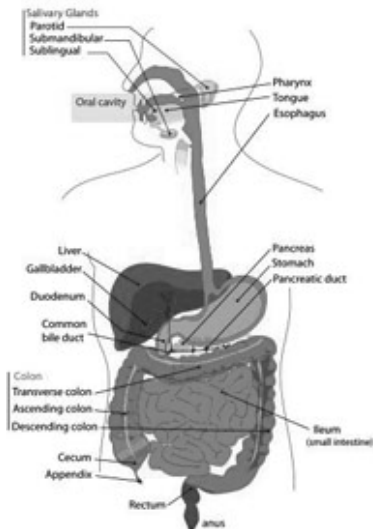


Fig. 2. *Upper alimentary tract*
(http://en.wikipedia.org/wiki/Upper_alimentary_tract, 15th March 2007)

- Image files are language independent, i.e. can be reused on sites that are written in different languages.
- Image files are independent from the wiki site, which means that contributors have the ability to donate or reuse the existing pictures.
- Contribution of images requires access to special equipment.
- Contributors need to respect international legal regulations, as copyright laws vary from country to country. Therefore, all licences of use must be explicitly specified by the contributors. A complete list of licences is usually provided on the home page <http://en.wikipedia.org/wiki/Wikipedia:Copyrights>.

All non-textual files, i.e. visual representations as well as sound and media files, are stored in the Wikipedia Commons repository. The Wikipedia Commons consti-



Fig. 3. *Uzdovské Peklany* (http://en.wikipedia.org/wiki/Uzdovské_Peklany, 15th March 2007

tutes a part of the Wikimedia Foundation which was designed with the aim of saving all the uploaded files for use in all the projects of Wikimedia. It is a central policy of the Wikipedia Commons that the files contributed to the project should be useful not only to the encyclopedia, but also to other projects (see Wikipedia Commons policy).¹⁸

The sources from which images are transferred to Wikipedia include mainly self-edited files constituting the work of the individual contributors themselves. The second major source of such files comprises public domains, i.e. governmental sites, university pages or society portals and book scans. The repository for August 2007 encompasses over 500,000 media files, categorised into different thematic sections. Though many of them are inserted into different entries of the encyclopedia as well as in other Wiki projects, many of the files stored in the repository will remain unused, being redundant or too personal for universal purposes, e.g. Fig. 4: p. 77.

The miscellany visible in the collection of the files results from the existence of a large number of contributors having different interests, passions and backgrounds.

Wikipedia also accepts the presentation of scientific formulae, presented in TeX markup language. All articles discussing problems concerning mathematical or physical sciences are accompanied by a visual presentation of the problem, an aspect not normally included in paper versions of general encyclopedias, due to the space restrictions mentioned above, e.g.

$$x_1 = \frac{b_1}{l_{1,1}}$$

$$x_2 = \frac{b_2 - l_{2,1}x_1}{l_{2,2}}$$

$$x_m = \frac{b_m - \sum_{i=1}^{m-1} l_{m,i}x_i}{l_{m,m}}$$

Upper triangular matrix (http://en.wikipedia.org/wiki/Upper_triangular_matrix, 18th March 2007)

¹⁸ http://commons.wikimedia.org/wiki/Commons:Community_Portal

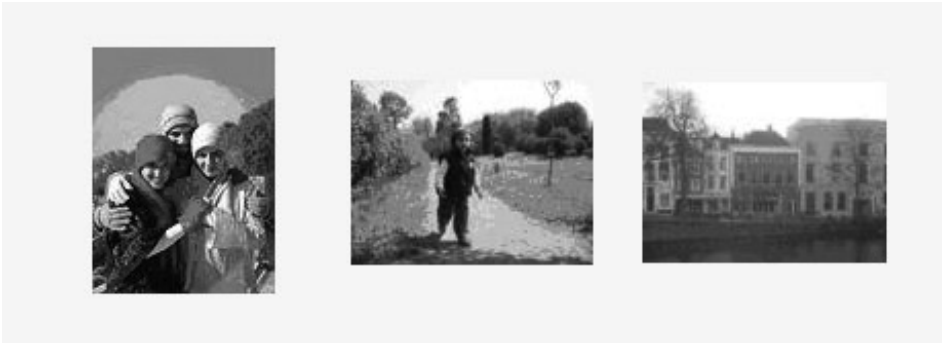


Fig. 4. (http://commons.wikimedia.org/wiki/Category:Pictures_and_images)

What is particularly attractive is the inclusion of animations presenting scientific phenomena. Such visual files undoubtedly facilitate the comprehension of a given problem, which frequently cannot be sufficiently described in a textual form. Examples of animations can be viewed at:

<http://commons.wikimedia.org/wiki/Category:Animations>.

A characteristic feature of Wikipedia, unlike traditional encyclopedias, is the presentation of the so-called “featured picture” – displayed every day on the front page of the encyclopedia, e.g.:



Fig. 5. A Monarch Butterfly (*Danaus plexippus*) on a Purple Coneflower (*Echinacea purpurea*)
Featured picture on 04. 10.2007 (http://en.wikipedia.org/wiki/Main_Page, 4th October 2007)

Featured pictures involve instances of the best work of the contributors. Materials proposed for this title are submitted to a review process before being awarded the status in question. The procedures for choosing the featured picture mainly take into consideration the degree of helpfulness of particular images in illustrating specific concepts.¹⁹

Moreover, in January 2008, the Wikimedia Foundation undertook an initiative aimed at the further enhancement of Wikipedia articles with rich-media content, i.e. with video films.²⁰ New software, which is currently being introduced and tested by users, allows the editing of images, sounds, animations and movies in a simple way, analogous to the editing of textual files. In a similar manner to the textual content, video files are created collaboratively, by the joint efforts of contributors.²¹ The incorporation of rich-media and the support of textual content with a visual demonstration will undoubtedly enhance the quality of the articles and facilitate the comprehension of the presented definitions. As the authors of the project claim, as a result of these files Wikipedia will “surpass any existing encyclopedia in the world.”²²

The multimedia of the free encyclopedia incorporate, apart from images, also audio files, since Wikipedia has launched a program for producing recordings of its articles. Until September 2007, nearly 800 articles have been recorded and are available online.²³ The authors of the encyclopedia are aware of the positive and negative sides of this undertaking and state them openly on the website. The advantages of creating spoken recordings are many: they make Wikipedia available to the visually-impaired (without the need for other appliances such as screen-readers) as well as to those who cannot read English, and can enhance the process of learning (provided the pronunciation is of high-quality). Though the initiative is praiseworthy and greatly expands the functionality level of the encyclopedia, it faces a few problematic issues. The most important of these, as Wikipedia authors state, is the continuous evolution of Wikipedia’s articles, which can make the recordings out-of-date and incompatible with the textual versions. Moreover, recording articles is time-consuming and requires accurate and impeccable pronunciation, so that it does not mislead non-native speakers.²⁴

Wikipedia is enlarging the scope of the multimedia on offer and, judging by the continuing development of the wiki-system, it is only a matter of time until more forms of novel data transmission will be added to the content.

¹⁹ As Viegas (2007) notes, the criteria for choosing featured articles seem arbitrary at times. Users admit that the dominance of photographs in the archives signifies that the “usefulness” criterion is not the standard for promotion. Otherwise, diagrams and maps would be equally represented.

²⁰ http://wikimediafoundation.org/wiki/Wikipedia_Invites_Users_to_Take_Part_in_Open%2C_Collaborative_Video_Experiment

²¹ http://wikimediafoundation.org/wiki/Collaborative_Video

²² http://www.kaltura.com/devwiki/index.php/Main_Page

²³ http://en.wikipedia.org/wiki/Category:Spoken_articles

²⁴ Ibid.

b) Bibliographic material

Wikipedia offers a range of different sorts of bibliographic material. It allows also the introduction into the entries of further links to online materials, thus providing direct access to the primary sources.

In order to secure a high quality of reference sources, the free encyclopedia provides guidelines and official policies, which set the standards to be followed by all users as far as the composition of bibliographic records are concerned. To the most important belong the policies of “Verifiability” and “No original research” which advise quoting only reviewed and established data (cf. Section 5.2.5). Moreover, contributors are required to consult multiple sources, preferably published peer-reviewed material, to omit dubious references and self-published web sites. Contributors are as well advised to use online databases, such as Google’s Book Search and the internet Archive’s Million Book Project, which provide access to full text sources in the case of books for which the copyright has expired.²⁵

The analysis of the material has proved that despite the above-mentioned policies advising adequate sourcing of the content, Wikipedia represents a considerable variation with regard to referencing, and many entries still lack proper sources. Out of 968 articles, 38% did not provide any references, attaching only links to external sources. Out of this group, 15% included an introductory remark stating that the following article did not contain proper references, and thus should not be fully trusted. Yet, there are still a large number of articles without references and without such a warning, which can mislead potential readers. The remaining group of 368 articles were characterized by a substantial multiplicity and heterogeneous scope of the references provided.

The following Tables present the percentage of references and the share of specific reference types within the analysed material:

Articles without reference	Articles with 1 reference	Articles with 2-20 references	Articles with 21-100 references	Articles with > 100 references	Articles with ‘citation needed’
38 %	12%	29%	9%	4%	8%

Table 3. Percentage of references

Book	Print article	Online article	Official web site	Private web site
62%	16%	11%	8%	3%

Table 4. Type of reference sources

²⁵ http://en.wikipedia.org/wiki/Wikipedia:Citing_sources

As shown in the Tables, the references include books, articles and web sites. The number of references differs considerably among the articles, ranging from one (*Ueuecoyotl*), through 49 (*Ubuntu*) or 69 (*Ubiquitin*) to 214 (*United States*). Interestingly, among the entries labelled “stubs”, i.e. short and incomplete articles, 29% provided a reference to at least one source.

Such percentage of references, though still insufficient, is higher than the number of references accompanying entries in *Britannica's Micropaedia*, where only 11 out of 535 articles contain lists of bibliographies, with the average number of sources amounting to 2 (highest 7, lowest 1). Wikipedia's sources may be even comparable to those provided by *Macropaedia*, which lists between 30 (*Uruguay*) and 379 (*Union of Soviet Republics*) bibliographic sources.

It is worth noting with regards to Wikipedia's sources the addition of an ISBN number of books in bibliography links, e.g.:

- Hill, Benjamin Mako. *The Official Ubuntu Book*. US: Prentice Hall. ISBN 0-13-243594-2 (*Ubuntu* <http://en.wikipedia.org/wiki/Ubuntu>, 6th April 2007),
- Toby A. H. Wilkinson, *Early Dynastic Egypt*, Routledge, London/New York 1999, ISBN 0-415-18633-1 (*Uenephes* <http://en.wikipedia.org/wiki/Uenephes>, 6th April 2007).

The addition of an ISBN number, however, has not yet become a common and normal practice among contributors, being introduced only into a small percentage of articles.

The material, though, abounds in entries (8%), which, instead of providing a proper reference, include a number of statements with a notation stating “citation needed” at the end, e.g.:

- Due to its presence in Group 11, it is a transition metal and so would probably appear as a heavy, solid, shiny metal.^[*citation needed*] (*Ununinium (Roentgenium)* <http://en.wikipedia.org/wiki/Ununinium>, 30th March 2007)
- William Thorburn is said to have been amazed by the city's beauty and hence settled here with his wife in 1822.^[*citation needed*] (*Uddevalla* <http://en.wikipedia.org/wiki/Uddevalla>, 6th April 2007))

In such instances, the authors introduce information of which they were not able to provide the source. This undermines the validity of data and raises doubts as far as its accuracy is concerned.

It is also worth mentioning that there are a number of high-quality, well-written articles, which did not provide any sources, such as: *Unitary matrix*, *Udden-Wentworth Scale*, *Uronic acid* and *Udeb* (file format). Apart from internal hypertextual links to other entries in the encyclopedia, these articles do not contain any further references to scholarly works.

The degree of referencing and the variability of further sources provided by Wikipedia is, additionally, not accompanied by a uniform presentation of such material, as there can be seen inconsistencies in the style of bibliographic description. Guidelines stating that references should follow the pattern: name of the author, title of the book, year of publication and name of the publisher, are frequently violated or ignored.

Inconstistency and chaos can be observed in the lack of a homogeneous notation of authors' names and titles.

- Karl Young, *The Continuum of Life in Codex Borbonicus (Ueuecoyotl* <http://en.wikipedia.org/wiki/Ueuecoyotl>, 6th April 2007)
- Turnbull, Stephen (1998), 'The Samurai Sourcebook'. London: Cassell&Co (*Uesugi Kagekatsu* http://en.wikipedia.org/wiki/Uesugi_Kagekatsu, 5th June 2007),
- HARRISON, PETER (1988): *Seabirds* (2nd ed.). Christopher Helm, London (*Uria* (<http://en.wikipedia.org/wiki/Uria>, 16th December 2007),
- John Abbatiello. *Anti-Sumbarine Warfare in World War I: British Naval Aviation and the Defeat of the U-Boats* (2005) (*U-boat* <http://en.wikipedia.org/wiki/U-boat>, 6th April 2007).

Such inconsistencies create the impression of the sloppiness and neglect of editorial guidelines.

5.2.2.4. Genre syncretism

A further feature of Wikipedia, contrary to its paper antecedents, is the tendency towards genre hybridization. Thus, apart from the encyclopedic content, Wikipedia includes the following genres:

- News – offering the “Latest news” section, thus approaching the function of a news portal.
- Curiosities – present in the “Did you know” section, which offers curiosities from the world of nature as well as rumours concerning the lives of personalities.
- Calendar – the encyclopedia offers an “On this day” section, where anniversaries or memories of specific historical events associated with a particular date are described.
- Discussion – Wikipedia offers hyperlinks to discussions concerning the project (see Wikipedia Community Portal) as well as to discussions concerning particular articles, where users can justify and debate the merits of their contributions. Thus, the structure of expression, which was monologic in traditional encyclopedias, can also take the form of a dialogue, or even a polilogue.
- Dictionary – the home page of Wikipedia offers links to a sister project, a Wikitionary, i.e. a dictionary and a thesaurus.
- Statistical resource – Wikipedia offers the content equivalent to a statistical yearbook, giving most recent data concerning population, based on public censuses or government reports.
- Yearbook – Wikipedia constitutes a source not only of current information, but also contains a summary of the events of a particular year across the world.
- Personal home-page – Wikipedia offers a section comprising User-portals, devoted to the contributors of the encyclopedia. These sites resemble personal web pages presenting the profiles of individual writers, their personal information as well as the history of their collaboration with Wikipedia.

5.2.3. Form

The transfer of the encyclopedia from print to the digital medium of the internet has triggered interesting changes in the form and overall organization of the encyclopedic content. The most meaningful transformations affect the density, methods of content classification and cross-referencing.

5.2.3.1. Density

Unlike traditional encyclopedias, the structure of the free encyclopedia is not driven by spatial restrictions and constitutes an example of a full exploitation of this “luxury of space” (Burke 2003: 246) afforded by the medium. Consequently, the structure of the entries has undergone considerable expansion. Each article in the free encyclopedia is presented as a separate and individual site. Besides, as Pang (1998) notices, authors, not being subject to word limits, can present indepth discussions of specific subjects. This privilege, however, if unreasonably exploited, may constitute a serious drawback (cf. Section 5.2.3.3.).

5.2.3.2. Article structure

The rise of the electronic media, as Görlach (2004: 180) points out, has initiated the modification of the so far relatively stable structure of entries, offering broader and more complex formats than those of printed works. The structure of Wikipedia entries is more flexible and unstable. Unlike in traditional encyclopedias, Wikipedia articles do not necessarily reach a final stage (unless blocked by administrators), as the content undergoes continual changes and revisions, and new additions are made instantaneously by successive authors.

An average Wikipedia article has the following structure:²⁶

²⁶ The enumerated component parts concern mainly longer articles. Shorter entries do not always have a table of contents or long lists of references.

Unicorn - Wikipedia, the free encyclopedia - Mozilla Firefox

PLK Edycja Widok Historia Załadki Narzędzia Pomoc

W http://en.wikipedia.org/wiki/Unicorn

Log in / create account

article discussion edit this page history

Unicorn

From Wikipedia, the free encyclopedia

For other uses, see *Unicorn* (disambiguation).

A **unicorn** (from Latin *unus* 'one' and *cornu* 'horn') is a *mythological creature*. Though the modern popular image of the unicorn is sometimes that of a horse differing only in the horn on its forehead, the traditional unicorn has a billy-goat beard, a lion's tail, and cloven hooves - these distinguish it from a horse.^[1] Marianna Mayer has observed (*The Unicorn and the Lake*), "The unicorn is the only fabulous beast that does not seem to have been conceived out of human fears. In even the earliest references he is fierce yet good, selfless yet solitary, but always mysteriously beautiful. He could be captured only by unfair means, and his single horn was said to neutralize poison."

Contents [hide]

- Unicorns in antiquity
- Medieval unicorns
- The hunt of the unicorn
- Heraldry
- Possible origins
 - Alleged skeletal evidence
 - Elasmotherium or rhinoceros
 - A single-horned goat
 - The kirin
 - The narwhal
 - The onyx
 - The eland

Alleged skeletal evidence

Among numerous finds of prehistoric bones found at *Einhornhöhle* (Unicorn Cave) in Germany's Harz Mountains, some were selected and reconstructed by the mayor, of Magdeburg, Otto von Guericke, as a unicorn in 1663. Claims that the so-called unicorn had only two legs (and was constructed from fossil bones of mammoths and other animals) are contradicted or explained by accounts that souvenir-seekers plundered the skeleton; these accounts further claim that, perhaps remarkably, the souvenir-hunters left the skull, with horn. The skeleton was examined by Gottfried Leibniz, who had previously doubted the existence of the unicorn, but was convinced thereby.^[12]

Baron Georges Cuvier maintained that as the unicorn was cloven-hoofed it must therefore have a cloven skull (making impossible the growth of a single horn); to disprove this, Dr. W. Franklin Dove, a University of Maine professor, artificially fused the horn buds of a calf together, creating a one-horned bull.^[14]

P. T. Barnum once exhibited a unicorn skeleton, which was exposed as a hoax.

Since the rhinoceros is the only known extant land animal to possess a single horn, it has often been supposed that the unicorn legend originated from encounters between Europeans and rhinoceroses. The Woolly Rhinoceros would have been quite familiar to ice age people, or the legend may have been based on the surviving rhinoceroses of Africa. Europeans and West Asians have visited Sub-Saharan Africa for as

See also

- Invisible Pink Unicorn (a modern satirical religious symbol)
- Shadhevar (a unicorn-like creature in Persian folklore)
- Horned Unicorn
- Qarnahuto (mythological bull-unicorn)

References

- Beer, Rüdiger Robert, *Unicorn: Myth and Reality* (1977). (Editions: ISBN 0-88405-583-3; ISBN 0-904069-15-X; ISBN 0-442-80583-7.)
- Encyclopædia Britannica*, 1911: "Unicorn"
- Gottfredsen, Lisa, *The Unicorn* (1999). (Editions: ISBN 0-7892-0595-5; ISBN 1-86046-267-7.)
- Shepard, Odel, *The Love of the Unicorn*. (1930) text
- The Living Unicorn

Notes

- ¹ Coincidentally, these modifications make the horned ungulate more realistic, since only cloven-hoofed animals have horns.
- ² Discussion of the Indus Valley Civilization with mention of unicorn seals
- ³ Site with slide show about unicorn seal
- ⁴ Jewish Encyclopedia
- ⁵ Ctesias (c.390 BC), "40", *India* (quoted by Photius)
- ⁶ Aristotle (c.350 BC), "3.2", *On the Parts of Animals*, trans. William Ogle.
- ⁷ Aristotle (c.340 BC), "2.1", *History of Animals*, trans. D'Arcy Wentworth Thompson.
- ⁸ Strabo (before 24 AD), "15.1.56", *Geography*.
- ⁹ Pliny (before 79 AD), "8.31", *Natural History*, trans. John Bostock. Also 8.30 and 11.105
- ¹⁰ A. Browne, Thomas (1646), "23 (Book 3)", *Pseudodoxia Epidemica*
- ¹¹ (Ashmolean Museum) "Young woman seated in a landscape with a unicorn", Leonardo, Late 1470s
- ¹² ²⁰⁰⁶ *file*, Stephen (1987), *A New Dictionary of Heraldry*, London: Alphabooks & C Black, p. 353-354, ISBN 0900670-446.
- ¹³ A Robin Meadows, "The Unicorn, the Mermaid, and the Centaur" Zoogooer, November-December 2006
- ¹⁴ A Dr Dove's Unicorn Bull. Retrieved on 2007-01-20.
- ¹⁵ A Man Made Unicorn. Retrieved on 2007-01-20.
- ¹⁶ A The Living Unicorn
- ¹⁷ A Unicorn at Gouttipedia. Retrieved on 2007-01-20.

External links

- All About Unicorns Historical unicorn information, plus a gallery of unicorn pictures

Wikimedia Commons has

Fig. 6. Wikipedia article structure (<http://en.wikipedia.org/wiki/Unicorn>, 18th March 2008)

Each Wikipedia article is accompanied by links to the following modes:

1. Editing mode – which enables a user to edit the content of the entry in order to introduce the desired modifications. The page assumes the following form in this mode:

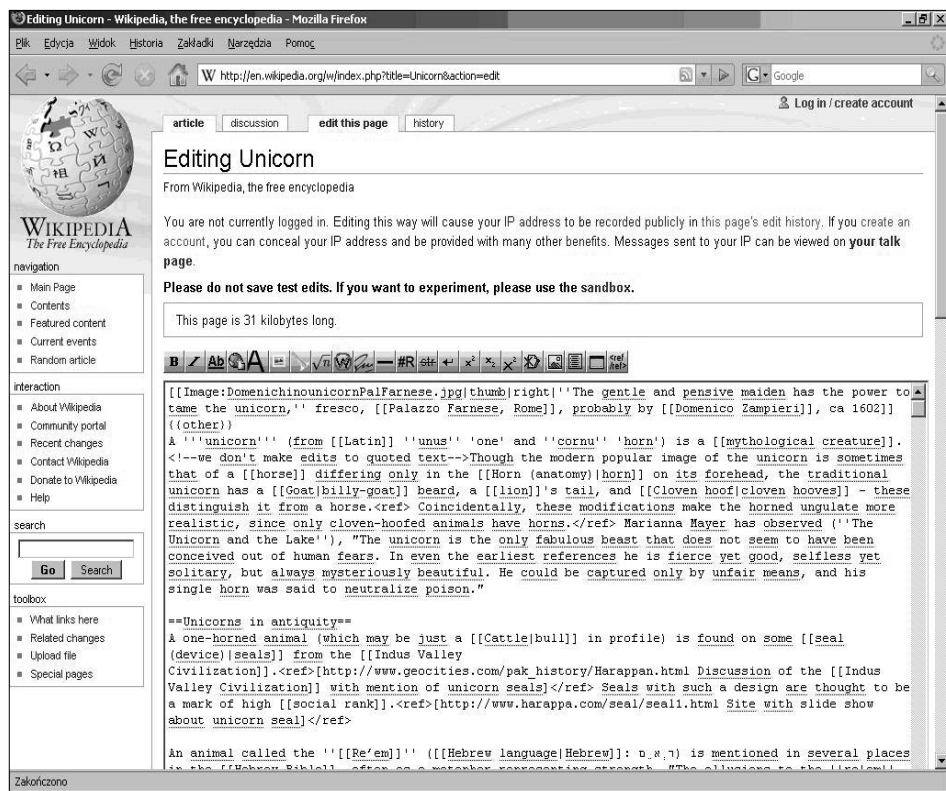


Fig. 7. An instance of an article in the editing mode (`http://en.wikipedia.org/w/index.php?title=Unicorn&action=edit`, 18th March 2008)

2. Discussion mode – this link directs the user to the talk pages concerning the article in question, where contributors post messages referring to the main article. A list of a few posts on the Talk page can be observed in the following example:

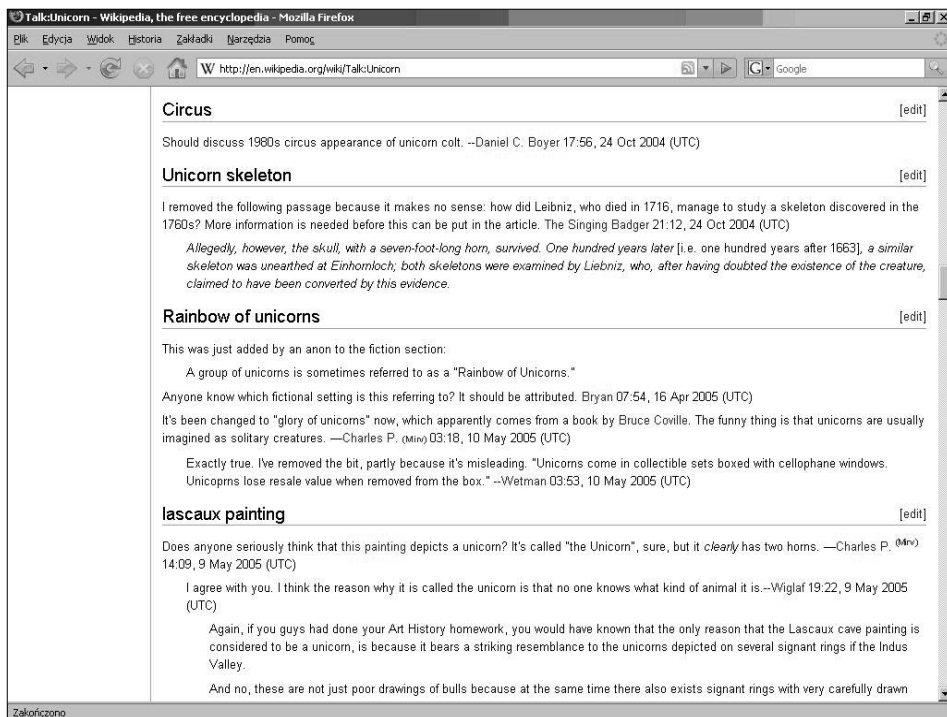


Fig. 8. An instance of a typical discussion page (<http://en.wikipedia.org/wiki/Talk:Unicorn>, 18th March 2008)

Stvilia *et al.* (2005) and Viegas *et al.* (2007) researched how the community of Wikipedians maintain and enhance the quality of information through discussions on talk pages. The studies confirm the significant role of these pages in the determination of the quality standards of Wikipedia articles, as it is here that the contributors debate on the final shape of an entry. The study conducted by Viegas *et al.* (*ibid.*) demonstrated that the posts contain suggestions for editing coordination, requests for information, references to vandalism, to Wikipedia guidelines, and voting sessions over controversial content and editing actions (*ibid.*: 7). It needs to be noted that talk pages are also places of heated disputes between editors disagreeing about the content of an article. Users as well often send aggressive or off-topic posts, varying in length and relevance (Sanger 2009).

3. History mode – where the history of the article is illustrated, with a list of all the contributions, user-names or IP addresses of the authors and a description of the changes introduced, e.g.:

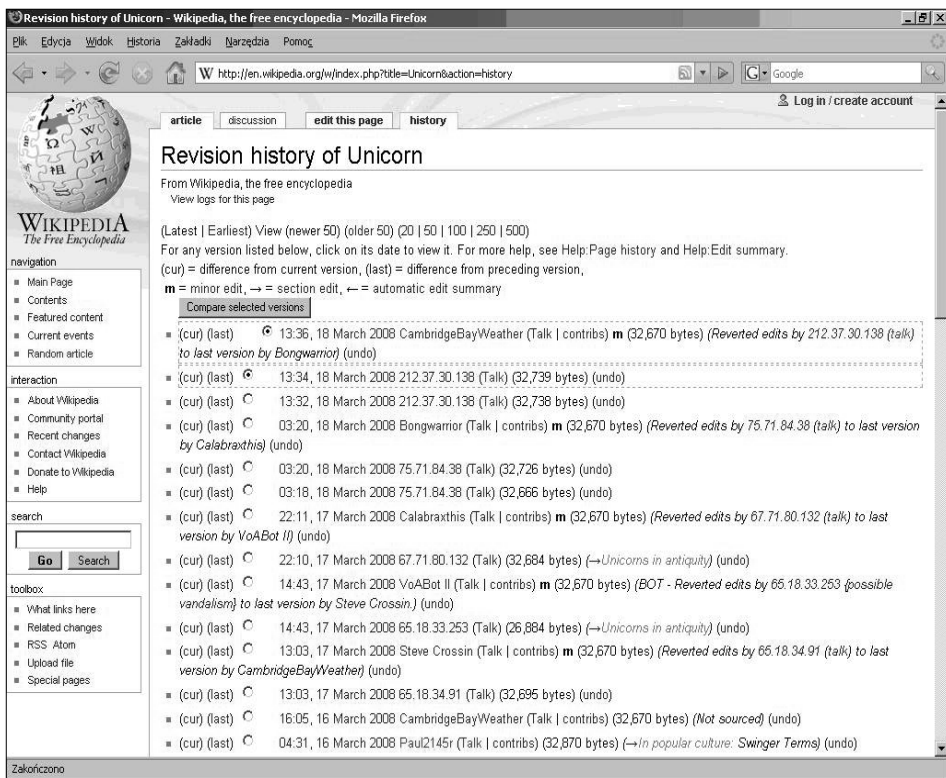


Fig. 9. An example of an article in the history mode (<http://en.wikipedia.org/w/index.php?title=Unicorn&action=history>, 18th March 2008)

This mode functions as an archive containing previous versions of the page and helps follow the modifications introduced and facilitates the reinstatement of the original version of an entry if the content has been vandalised or deleted.

Wikipedia articles are varied with regard to their length and content.²⁷ Often very extensive and thorough, the entries reach beyond the basic encyclopedic knowledge, more closely resembling scientific articles. The analysed material consisting of 968 entries included articles of both one sentence and several page in length. The mean number of words per article amounted to 800, with the least being 7 (*Upper Knowle*) and the largest 11,216 (*United States*). The length of articles is slightly higher than the entries found in *Micropaedia*, where the average number of words per article is 650. The articles are shorter than the entries offered by *Macropaedia*, all above 1,000 words, with the average number being 7000 words per article, covering from 6 (*Uruguay*) to 139 two-column pages (*United Kingdom*). This difference, however, occurs due to technological reasons, since longer articles in Wikipedia tend to be di-

²⁷ The variation also concerns referencing, which will be discussed below, and illustrating, indicated above.

vided into shorter entries, all of them connected by hyperlinks. This is done also for the sake of the convenience of reading, as it is easier to process and follow an entry which is one or two screen-pages long at the most.

Detailed data concerning the length of the reviewed articles are presented in the following Table:

One sentence articles	10–200 words articles	201–1000 words articles	> 1000 words articles	> 10 000 words articles
98	152	484	216	18

Table 5. Length of articles in Wikipedia

In the case of longer articles, graphic means of segmentation play a significant role. Spatial ordering of articles involves the use of tables of contents, subtitles and the differentiation of types of font sizes. The majority of the articles are characterized by a clear layout and a harmonious composition of textual and graphic elements. Yet, there is a certain inconsistency visible in the use of content tables. In the case of long articles, they constitute important guidelines and indicators of the information covered by a particular entry. However, within the analysed material, there occurred on a few occasions the unnecessary introduction of a list of contents (e.g. *Uberlandia*, *Uccello*, *Usnwc*). Due to the limited length of these articles and the meagre variety of data provided by the entry, the use of a table of contents seems redundant. Similarly, this irregularity in the use of lists of contents can be illustrated by the articles on *Udo Bolts* and *Udo Jurgens*, both of comparable length and content, yet only the latter containing a table of contents.

These inconsistencies represent an obvious consequence of the freedom of contribution offered by Wikipedia, where the construction of an entry rests entirely with the author.

5.2.3.3. Classification

The electronic form enables nearly all possible ways of cataloguing content, and thus facilitates the organization of large amounts of data. As Bolter (2001: 88) observes, in an electronic environment, the alphabetical categorization does no longer constitute the primary order of the text. Due to the affordances of technology, as the scholar claims, the topical classification has regained an important status among encyclopedic publications. Similarly, McArthur (1986: 182) claims: “in the electronic culture of the late 20th century people are acclimatizing quickly to thematic order and complex cross-referencing. In the process, these devices are losing their low peripheral status and in tandem with the PC are gaining high cultural value. In effect, the pendulum is swinging back again from an era in which alphabetic order attained high status in print culture.”

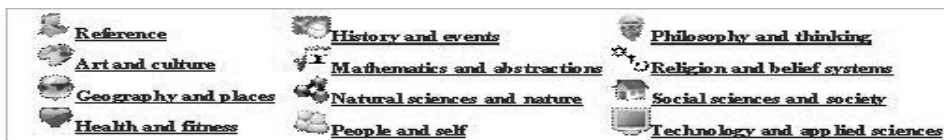


Fig. 10. List of Wikipedia's categories (http://en.wikipedia.org/wiki/Portal:Contents/Categorical_index, 8th March 2007))

Moreover, the electronic form allows the combination of different patterns of classification which can coexist within one interface.²⁸ The organization of the content depends, we might say, purely on the editors' and users' preferences.²⁹

The free encyclopedia offers three broad ways of content listing, which include indices, lists and featured content.

I. Indices include the standard, traditional forms of content organization:

1. Alphabetical index – presenting all articles in alphabetical order, sorted by the first two letters of the title, e.g.: *Fa Fb Fc* as well as capital letters *FA FB FC* for acronyms or contractions.³⁰
2. Categorical index – which presents an index of twelve main thematic categories of articles referring to different disciplines of knowledge, further subdivided into detailed sections.³¹

Within the categorical index, Wikipedia offers sub-categorizations into three worldwide accepted forms of classification, i.e. Dewey's classes (listing two levels of this system), the Library of Congress classes, and Roget's Thesaurus (with articles arranged according to six classes plus further branches).

II. Lists – which include several types:

1. Overviews – presenting key articles arranged by subject and also lists of Wikipedia's articles divided according to academic disciplines.
2. Lists of topics – involving lists of key topics as well as two broad collections comprising lists of countries and lists of people.

²⁸ It has been observed (Katz 1998, Bolter 2001) that the 15th edition of the *Encyclopedia Britannica*, due to its construction, could be regarded as a forerunner of the electronically fashioned encyclopedia. Its arrangement and the system of references, hard to follow in a printed book, would constitute the perfect organization of content in an electronic text, where hyperlinks could facilitate the process of moving across the outline and the main body of articles.

²⁹ However, as Bolter (2001: 88) notices: "any single topical outline today must seem arbitrary because it reflects one editorial view of the organization of knowledge, which the reader may not share or even comprehend." The scholar quotes the words of D'Alembert uttered when he compared the encyclopedia to a world map and said that "...one can create as many different systems of human knowledge as there are world maps having different projections. There are hardly any scholars who do not readily assume that their own science is at the centre of all the rest, somewhat in the way that the first men placed themselves at the centre of the world."

³⁰ http://en.wikipedia.org/wiki/Portal:Contents/Quick_index, 19th March 2007.

³¹ http://en.wikipedia.org/wiki/Portal:Contents/Categorical_index, 19th March 2007.

3. Portals – described as introductory pages; they include selected articles, pictures and news on a particular subject, presenting further information on the topic and gathering users interested in the subject in question. Portals offered by Wikipedia include the following: Arts and Culture, Current Events, Geography, Health, History, Natural Sciences and Mathematics, Philosophy, Religion and Spirituality, Social Sciences and Society, Technology and Engineering.
4. Glossaries – presenting lists of definitions of specialized words. Glossaries are arranged by subject and also alphabetically into categories.
5. Timelines – involving lists of articles arranged chronologically. Timelines include also lists of centuries, decades, historical anniversaries, current events and graphic timelines arranged alphabetically.³²

III. Featured content – involving the best materials stored in the encyclopedia. Featured materials are divided into textual, graphic and audio groups and those presenting the best portals and topics.³³

Categories and lists of the free encyclopedia form a complex network of interconnected sections. Due to the simultaneous coexistence of multiple-categorization schemes each article can belong to several categories and be included in a portal or overview.

The digital format of the encyclopedia resolves as well the problem of finding appropriate information with the introduction of automatic searching, a valuable facility which was not available in traditional encyclopedias. The simplest procedure to find an entry, for those who know the name of the article, is to type the term into the search box. Upon pressing the “Go” button, the system automatically directs users to the specific article. An important feature facilitating access to a desired entry is the fact that users do not need to know the exact spelling of an entry word. Wikipedia has developed a tool which tries to re-route the reader to the desired data by presenting lists of search results which are associated with the term typed in the search box. The system presents the reader with a list of articles thematically associated with the term, together with the approximate relevance of the result to the desired content.³⁴ It is also possible to limit search options by enclosing the search phrase into quotation marks, which restricts results to an exact finding.

Moreover, Wikipedia offers a “random choice” search facility, which provides a user not looking for any specific data with entries randomly chosen from the contents of the whole encyclopedia.

The continuously increasing number of entries might, however, pose serious problems in the future as far as the content classification is concerned. Storage capacities have risen, but methods to avoid a deluge of information are still much underdeveloped. As Hanks (2001: 9) rightly points out, “the absence of space constraints calls for more, not less intellectual discipline in the selection and arrangement of information.”

³² <http://en.wikipedia.org/wiki/Portal:Contents>, 19th March 2007.

³³ http://en.wikipedia.org/wiki/Portal:Featured_content, 19th March 2007)

³⁴ For instance: search results for “Joao Cabral” give us the following: Joao Cabral de Melo Neto (100%); Joao Bosco Cabral (96.5%); Colonial Heads of Sao Tome and Principe (73.6%); Luis Cabral (71.9%), etc.

The first signals of information overload are already visible in the English edition of Wikipedia. For example: a user searching for information on “John Smith”, upon entering the name in the search box, is given the disambiguation list containing 87 articles entitled *John Smith*.³⁵ Finding the appropriate one necessitates reviewing the whole list. Therefore, it would seem useful to design an improved and more effective system of classification and disambiguation based, for instance, on a detailed system of categories.

5.2.3.4. Cross-references

A further property of Wikipedia, absent from traditional print encyclopedias, is a highly complex system of hyperlinks. Hyperlinks are highlighted elements of hypertexts, which upon clicking lead the reader to a different part of the text or to a different document. The aim of the hyperlinks in Wikipedia is to direct users to further materials on the subject in question. Hyperlinks not only increase the amount of information provided to the reader, as compared with a print source, but also facilitate access to information. There are two main groups of links: internal and external.

a) Internal links

Internal links refer to hyperlinks within the free encyclopedia. They comprise hyperlinks within an article, where the constituent words lead readers to other articles in Wikipedia associated with the specific word (these articles may not contain any materials associated with the entry in question). Wikipedia is characterized by a high density of internal links. Within the analysed material, there did not occur any article without hyperlinks, with the smallest number of links being one (*Upper Knowles*), and the largest 1,179 (*United States*).

Apart from these links, Wikipedia offers also links to other materials thematically connected with the subject. These links are placed under the heading “See also”.

For instance: “See also” links to *Unified Energy System* (http://en.wikipedia.org/wiki/RAO_UES, 25th March 2007):

- Energy policy of Russia,
- Third Generation Company of Electric Power Wholesalers.

Within the present study, 47% of the articles provided this type of link and the mean number of such references amounted to five links.

b) External links

External links direct readers towards further materials associated with the subject, but these are found outside the free encyclopedia and include primary sources or additional materials.

For instance: External links for *U-Boat* (<http://en.wikipedia.org/wiki/U-boat>, 6th April 2007):

- www.uboaat.net,
- www.karldoenitz.com,
- Book excerpt: U-Boat attack,
- Wrecks.

³⁵ http://en.wikipedia.org/wiki/John_Smith_%28disambiguation%29

Within these links, online sources are clearly favoured among contributors, as nearly 60% (56.3%) of the external links comprise online materials. These were either links to pages presenting the content referring to the issue covered by the article, or links to open access research, i.e. peer-reviewed journal articles that have been made freely available online, e.g.:

- Kirkpatrick et al, J. Davy (July 10, 1999). “Dwarfs Cooler than M: the Definition of Spectral Type L Using Discovery from the 2- μ ALL-SKY Survey (2MASS)”. *Astrophysical Journal* 519 (2): 802–833 (Ubv System, http://en.wikipedia.org/wiki/Ubv_system, 6th April 2007),
- Peters JM, Franke WW, Kleinschmidt JA. (1994) Distinct 19S and 20S subcomplexes of the 26S proteasome and their distribution in the nucleus and the cytoplasm. *J Biol Chem*, March 11;269(10):7709–18. PMID 8125997 (*Proteasome* http://en.wikipedia.org/wiki/Ubiquitin-proteasome_pathway, 6th April 2007).

The addition of links to the studies available online, which can be easily consulted by the readers, considerably increases the authority, reliability and educational quality of the encyclopedia. As Willinsky (2007) claims, this capacity of online access to other resources could be further utilized by Wikipedia to encompass the provision of links to journals accessible online dealing with particular subject, so that readers would be able to proceed immediately to the latest reliable information which is available.

However, in Wikipedia the use of external links is at times exaggerated, as some links are introduced due to the personal interests of the authors or verge on the unsuitable or inappropriate. For instance:

- Udo Jurgens Fan Page (http://en.wikipedia.org/wiki/Udo_Jurgens, 6th April 2007),
- Photo of Uday’s Corpse (Warning) (http://en.wikipedia.org/wiki/Uday_Hussein, 29th November 2007).

Moreover, as Wallace and Van Fleet (2005: 102) observe, the density of links may distract the reader and turn him/her aside from the main purpose of consulting the work.

Hyperlinking has exerted considerable influence on the composition of the articles. As Pang (1998: para 8) notices: “print encouraged editors to think about articles as self contained objects, but in the electronic world the article is starting to become as obsolete a unit of editorial planning and work as it is becoming irrelevant to readers.” In the electronic media, as the scholar observes, articles form one huge body of information, unlike in print, where the entries to an encyclopedia exist as disparate individual objects. Pang (*ibid.*) further claims that this fact forces upon editors a different approach to an encyclopedia, i.e. they need to perceive it as an interconnected whole, where the smallest changes to one entry may affect the whole content. Moreover, hyperlinking between articles has initiated the process of hyperreading, where readers approach texts in a nonlinear manner, moving freely through different interconnected data. An author of a hypertext document should be aware of the specific rules guiding the creation of an electronic text. This is especially important as far as an encyclopedic article is concerned, frequently approached by readers not knowledgeable in a subject. Therefore, it is the contributors’ role to create an entry

which will provide users with appropriate data by introducing hyperlinks to materials which are relevant to the understanding of the subject.

Considering the affordances awarded by technology, the free encyclopedia has the potential to become a rich source of references.

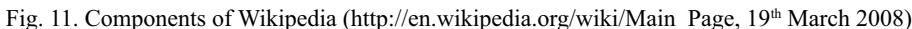
5.2.4. Functionality

What substantially differentiates Wikipedia from its paper antecedent is the functionality afforded by the medium. The transfer of the genre to the medium of the internet initiated far-reaching changes in the services provided by the site. The functionality and interactivity options offered by Wikipedia include the following:

- Searching and browsing through the content – Wikipedia offers different ways of browsing the content, connected with various cataloguing systems, illustrated above. Browsing through the encyclopedia is facilitated also due to the above-mentioned high level of hyperlinking.
- Downloading the data – users can download the entire content of Wikipedia (version Wikipedia 1.0), with the additional possibility of obtaining systematic updates. Users can also download parts of the encyclopedia, entries, articles or reference lists of particular interest.
- Constant updates – Wikipedia, unlike paper encyclopedias, can remain up-to-date with latest news, as the system allows instantaneous addition and modification of the contents.³⁶
- Login function – users can create an individual user account within the site. This gives them access to the background of the encyclopedia, enables the creation of new pages, page renaming and image uploading as well as eligibility to vote in the Committee and Board elections and to customize the appearance of the web site.
- Discussions online – Wikipedia offers its users the opportunity to conduct discussions concerning the contents of the encyclopedia.
- Reference desk online – the site provides an opportunity for users to post questions or requests for an explanation of a problematic issue. The questions are answered by other users of the encyclopedia.
- E-mail option – Wikipedia provides options for the e-mail contact. This enables asynchronous communication with the creators of the encyclopedia.
- Contribution – as described above, the possibility of submitting contributions to the encyclopedia by all the users of the internet constitutes the most important characteristic of the encyclopedia. This fact is central to the uniqueness of Wikipedia as a genre of encyclopedia and constitutes its essence and determines its character.

The hybrid structure of Wikipedia is illustrated in Fig. 11:

³⁶ For instance the information about granting this year's Noble Prize Winner in Literature to Doris Lessing was published in Wikipedia a minute after the Jury's verdict (official statement was issued at 11.00 on 11.10.2007 and introduced to Wikipedia on the same day at 11.01, which can be seen in the history of the article http://en.wikipedia.org/wiki/Doris_Lessing).



5.2.5. Meeting the quality standards

In the previous chapter we referred to accuracy, validity and currency as the dominant characteristics with which a typical encyclopedia is commonly associated.

The currency of Wikipedia is achieved by the aforementioned system of continuous up-dating. Its compliance with the principles of accuracy and validity, though, constitutes a disputable issue. The quality of the entries is considerably uneven, with some of them being instances of exquisite research and topic presentation and others of low value, being incomplete and inaccurate.

The reliability and quality standards of Wikipedia have been repeatedly questioned by different authorities, scholars and journalists (e.g. McHenry 2004, Orłowski 2005, Gann 2007, Keen 2007). It has been observed that, though articles may be stylistically correct and may seem trustworthy, they may contain obscure and false information (Copestake 2006).³⁷ The tests of Wikipedia's reliability carried out by the science journal *Nature* in December 2005 showed that the free encyclopedia contained one-third more mistakes than the *Encyclopedia Britannica*.³⁸ As the main reasons for these errors, Halloway *et al.* (2005: 4) identify the lack of contributions from authorities in particular fields of knowledge as well as the lack of formal peer-review process which would verify the data. As a result, Wikipedia does not guarantee authenticity, which constitutes the main differentiating factor between Wikipedia and a standard encyclopedia and inhibits it from representing a proper reference source.

Additionally, a number of other factors do not let Wikipedia achieve the standards of an encyclopedia and constitute reasons for the low quality of the articles, e.g.:

- Malleability and the ephemeral nature of the content – due to the continuous updates, the content of the encyclopedia constantly evolves. As Mainguy (2006) observes, these constant changes within Wikipedia prevent its use for scholarly research, even when articles are comprehensive and correct.³⁹
- Vulnerability to ignorance and vandalism – the popularity of Wikipedia as well as its structure and open-access lead to numerous acts of vandalism concerning the content of the encyclopedia, which, even if corrected within a limited time-span, may considerably undermine the accurateness of the encyclopedia.⁴⁰

³⁷ Instances of erroneous data are described in McHenry (2004, 2005), Rosenzweig (2006).

³⁸ http://www.theregister.co.uk/2005/12/16/wikipedia_britannica_science_comparison. It needs to be noted that the *Encyclopedia Britannica* refuted the results of this test, pointing out many inconsistencies and fallacies in the procedures of the study.

³⁹ Relevant information could be retrieved from the archives and history pages. Such an approach is, however, cumbersome and impractical in its application.

⁴⁰ For instance, in July 2005 a fallacious entry reporting the death of a comic strip author, Jeph Jacques was available on Wikipedia for two days, whereas a false article about John Seigenthaler stating that he was directly involved in the Kennedy assassination remained unchanged for a month: http://en.wikipedia.org/w/index.php?title=Seigenthaler_controversy&oldid=169194180

- Systemic bias – connected with the personal characteristics and preferences of the contributors. Wikipedia, namely, tends to favour topics of interest to the largest groups of authors (Halloway *et al.* 2005: 4).⁴¹
- A lack of expert-friendly policies – as Sanger (2009) claims, the lack of expert recognition weakens the quality of articles and allows unknowledgeable editors to force their opinions.

In order to become an authoritative reference point, Wikipedia needs to overcome the above-mentioned hindrances. The first step that the community of Wikipedians has taken is the introduction of a set of protective mechanisms, deployed with the aim of ensuring generic integrity, i.e. the degree of an article's compliance with the conventions. The most important mechanisms include the following:⁴²

- Neutral point of view – this policy advises against the presentation of personal, subjective opinions. The content should not be biased, editors should not impose their personal viewpoint on the readers.
- Verifiability – the policy in this regard plainly states the following: “the threshold for inclusion in Wikipedia is verifiability, not truth.” “Verifiable” denotes data which have been published elsewhere. The policy continues: “facts, viewpoints, theories and arguments may only be included in articles if they have already been published by reliable and reputable sources.”
- No original research – in order to enhance accuracy and ensure reliability, Wikipedia advises against the usage of “original research”, i.e. research which has not yet been properly verified and approved as well as original thoughts or ideas which have not been recognized.
- Biographies of living persons – the policy advises against the presentation of controversial and negative information about living persons, without verifiable and reliable facts.
- User-identification – users are asked to share their identity (anonymous users, as mentioned above, are identified and monitored by an IP tracking procedure). Users who do not observe this regulation may be banned and blocked from accessing the contents.
- Awarding of the “protected” status to a site – articles presenting controversial or “sensitive” issues can be partially blocked by the administrators. Such articles can be changed only by users registered for a certain length of time.⁴³
- Reviewing procedures – volunteers and administrators review the suitability of new articles or the modifications introduced to the already existing entries. Articles which have been found to contain inaccuracies or which do not fulfil the basic requirements outlined in the guidelines provided by the free encyclopedia are assigned labels which function as a warning to the readers that the particular

⁴¹ Wikipedia reports imbalanced coverage by geographic area favouring North America, Japan, Western Europe, Australia and New Zealand, which are significantly more pronounced than in editorially created encyclopedias.

⁴² http://en.wikipedia.org/wiki/Wikipedia:List_of_policies

⁴³ The “protected” status has been for instance awarded to entries on George W. Bush, War in Iraq, or Islam.

article does not comply with the standard properties of an encyclopedic article. The warnings include the following:

This article may require cleanup to meet Wikipedia's quality standards.
Please improve this article if you can.

The neutrality of this article is disputed.
Please do not remove this message until the dispute is resolved.

This article needs additional references or sources for verification.
Please help improve this article by adding reliable references.

This article or section may contain original research or unattributed claims.
Please help Wikipedia by adding references.

Some information in this article or section is not attributed to sources and may not be reliable. Please check for inaccuracies and modify and cite sources as needed.

Specifically, as the labels illustrate, it is the articles which violate the condition of accuracy, neutrality, verifiability and suitable referencing that Wiki-authors treat as non-encyclopedic.⁴⁴ However, the collected material of 968 articles confirmed that the above enumerated labels are assigned selectively, i.e. not all the articles violating the principles of composition carry such a warning. Within the analysed set of entries, only 2.68% (26 entries) were assigned such a warning tag, whereas violations of the aforementioned principles were present in a considerably higher number of entries.

- “Featured article” status – in 2004 Wikipedia has introduced a reviewing system of the so-called “featured articles”, i.e. articles which are assessed for validity, completeness, objectivity and style, and comprise the group of premium entries.
- Fact-checking procedures – Wikipedia has also introduced a project centered on checking facts and references using multiple sources.⁴⁵
- Recent changes mode – each node in the encyclopedia has a log of all contributions and changes made to the content, together with the name or IP address of the contributor. This is helpful not only to identify vandals, but also to record the contribution actions of the registered users.

These mechanisms are aimed at giving Wikipedia a format which complies with the established standards of an encyclopedic work. Still, vandalism and inaccuracies continue to occur, which constitutes one of the most consequential disadvantages of the project. As Rosenzweig (2006) rightly observes, since Wikipedia is the most popular reference work on the web and is available under the GNU license, false information and incorrect definitions may be easily propagated on a wide scale to millions of its users.

⁴⁴ The form of these labels signifies the deeply interactive and dialogical character of the free encyclopedia. Readers are directly encouraged to co-operate in the re-production and improvement of the existing articles.

⁴⁵ http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Fact_and_Reference_Check

In conclusion, the comparison of the structure of Wikipedia with a traditional print encyclopedia shows that the two genres differ in many significant areas of genological characterization. The most important modifications comprise the extension of the function of the encyclopedia, editorial roles of the writer and the reader, content enhancement and structural alterations. Interestingly, Wikipedia exemplifies not only an enhancement of genre properties, but also a departure from the characteristic features conventionally associated with the genre, i.e. accuracy and validity of the content. The editorial freedom constitutes as well the basic cause of the irregularities evident in all areas of the free encyclopedia. These inconsistencies result both from contributors' ignorance of the existing rules and from the lack of a centralized enforcement of compliance with the guidelines.

Departure from the convention is also noticeable in discourse properties of Wikipedia articles, illustrated in the next part of the publication.

6. Wikipedia discourse features

6.1. The continuum of Wikientries

As mentioned above, the reproduction of a specific genre depends to a large extent on the knowledge of genre conventions and on an author's linguistic awareness, but also on the medium in which the genre appears. As Yates and Orlikowski (1992) observe, the nature of the medium, its affordances and constraints, guide users towards specific communicative choices. Wikipedia constitutes the product of a special environment, which is digital, open and collaborative. Therefore, it is relevant to investigate the influence of this environment on the discourse structure of the free encyclopedia.

Since Wikipedia is created collectively by volunteers, it is interesting to study the process of choice the authors make between traditional discourse structures and conventional language means and the introduction of generically foreign elements (i.e. going beyond the convention) as well as to study the influence of communicative competence on individual ways of editing. A further aim of the analysis is to verify our initial hypothesis concerning the discourse layer of Wikipedia and to find out whether articles are structured according to recurring principles, i.e. whether the authors apply homogeneous norms and rules when creating articles as well as to assess the influence of authors' individual stylistic choices on the realization of the genre and, therefore, on its changes.

As stated in the introduction, the analysis covers all the entries beginning with letter *U* available in Wikipedia, that is 968 articles.¹ The entries were analysed at different levels with respect to dominant discourse properties. Discourse features which are prototypically associated with an encyclopedic article were identified as well as the properties which depart from such conventions.

Clusters of discourse features characterizing Wikipedia entries allow us to extract three main groups of articles characterized by distinct properties. I propose to arrange the entries along a continuum reflecting the degree of their compliance with the principles defining the conventional properties of encyclopedic entries enumerated

¹ All the quoted articles are presented in their original form, maintaining original spelling and punctuation.

in Chapter III Section 3.2.3.5. Thus, at one end of the spectrum, we propose to place encyclopedic articles exhibiting features conventionally associated with the discourse of an encyclopedia and, at the other, anti-encyclopedic articles which exhibit far-reaching violations of the genre, with a large group of articles located between the two poles, i.e. unconventional articles, manifesting different degrees of modification of the genre:

The Continuum of Wiki-entries

Conventional ←-----→ **Unconventional** ←-----→ **Anti-encyclopedic**

It needs to be noted, though, that the boundary line between particular groups of articles is not definite and fixed. Wikipedia is varied and heterogeneous to the extent that the features characterizing a particular group of articles along the continuum may with differing intensity occur in the entries placed at the opposite pole of the dimension.

Each group of entries is described individually in the subsequent sections. The analysis presents the dominant properties of the respective groups of Wiki-entries. Since the group of unconventional articles constitutes the largest collection of entries, it is given a more extensive description which covers textual patterning as well as syntactic and lexical discourse properties of these articles. The analysis is predominantly qualitative in character, yet quantitative data is included in order to provide a statistical illustration of the occurrence of a given discourse feature to highlight the scope of a particular phenomenon.

6.1.1. Conventional articles

Nearly 23% (219 entries – 22.62%) of the articles from the entire analysed collection comprising 968 entries are characterized by a close accordance with the requirements defining a prototypical encyclopedic article. These articles are compiled in line with the conventional properties associated with an entry in a traditional encyclopedia.

As such, these articles exhibit a high level of formality, achieved by the use of standard, complex syntactic structures and formal lexical means. In addition, such entries demonstrate a high degree of logical consistency, coherence and a clear structure.

Articles included in this group discuss different issues, ranging from medicine, history, through literature to politics. The length of these articles oscillates between 280 and 9840 words.

To present an exemplification of the compliance these articles exhibit with conventional encyclopedic articles, we present the juxtaposition of two extracts from Wikipedia articles with their thematic equivalents from the *Encyclopedia Britannica*. In order to provide a wider perspective for the following comparison, the chosen pairs of articles concern different domains of knowledge and represent different degrees of structural complexity.

The first pair of articles comprises the entries on *Uvarovite*, a chemical compound associated with the scientific domain of mineralogy. In both encyclopedias, the entries discussing this topic have a short, concise form:

Encyclopedia Britannica

Uvarovite - calcium chromium garnet found as small, brilliant, green crystals. It is the rarest of all the garnets, and its crystals commonly are too small to be cut. Otherwise, it would rival emerald as a popular gemstone because of its beautiful colour. Typical occurrences are in chromite, as in the northern Urals, California, Canada, Finland, and Silesia in Poland (1990, *vol.* 12: 224).

Wikipedia

Uvarovite is a chromium bearing garnet group species with the formula: $\text{Ca}_3\text{Cr}_2(\text{SiO}_4)_3$. It was discovered in 1832 by Germain Henri Hess who named it after Count Sergei Semenovitch Uvarov (1765-1855), a Russian statesman and amateur mineral collector.

Uvarovite is one of the rarer of the garnet group minerals, and is the only consistently green garnet species, with a beautiful emerald-green color. It occurs as well-formed fine sized crystals. Specimens of uvarovite are much sought after by collectors for outstanding brilliance and color. It is found associated with chromium ores in Spain, Russia, and Quebec in Canada. It also occurs in Finland, Norway, and South Africa. (<http://en.wikipedia.org/wiki/Uvarovite>, 24th January 2007)

The second pair of articles discusses a subject from the field of humanities and deals with *Ukiyo-e*, a genre of Japanese artwork. For the sake of space, we present only fragments of these articles, as in both encyclopedias they are quite lengthy:

Encyclopedia Britannica

Ukiyo-e: (Japanese: “pictures of the floating world”), one of the most important genres of art of the Tokugawa period (1603–1867) in Japan. The style is a mixture of the realistic narrative of the emaki (“picture scrolls”) produced in the Kamakura period and the mature decorative style of the Momoyama and Tokugawa periods. The ukiyo-e style also has about it something of both native and foreign realism. Screen paintings were the first works to be done in the style. These depicted aspects of the entertainment quarters of Edo (modern Tokyo) and other urban centres.

Hishikawa Moronobu is generally accredited as the first master of ukiyo-e. The transition from single- to two-colour prints was made by <http://www.britannica.com/eb/to pic?idxStructId=426612&typeId=13> Okumura Masanobu. In 1765 polychrome prints using numerous blocks were introduced by <http://www.britannica.com/eb/topic?idxStructId=575903&typeId=13> Suzuki Harunobu. The essence of the ukiyo-e style was embodied in the works of Utamaro, Hokusai, and Hiroshige (*ibid.*:110).

Wikipedia

Ukiyo-e “pictures of the floating world”, is a genre of Japanese woodblock prints (or woodcuts) and paintings produced between the 17th and the 20th centuries, featuring motifs of landscapes, tales from history, the theatre and pleasure quarters. It is the main artistic genre of woodblock printing in Japan.

Ukiyo, meaning „floating world”, refers to the impetuous young culture that bloomed in the urban centers of Edo (modern-day Tokyo), Osaka, and Kyoto that were a world unto themselves. It is an ironic allusion to the homophone term “Sorrowful World”, the earthly plane of death and rebirth from which Buddhists sought release. The art form rose to great popularity in the metropolitan culture of Edo (Tokyo) during the second half of the 17th century, originating with the single-color works of Hishikawa Moronobu in the 1670s. At first, only India ink was used, then some prints were manually colored with a brush, but in the 18th century Suzuki Harunobu developed the technique of polychrome printing to produce nishiki-e.
(<http://en.wikipedia.org/wiki/Ukiyo-e>, 4th January 2007)

As the examples show, there is an evident similarity between the two article pairs concerning their textual, syntactic and lexical structures.

Both pairs of articles exhibit an adequate level of clarity, achieved by a coherent presentation of successive information concerning the topics in question. This greatly enhances the level of readability and comprehensibility of the entries. Moreover, the articles reveal a high degree of density and conciseness, which adds to their formal character. Yet, despite their compact nature, both entries represent an analogous level of informativity.

The articles from both sources demonstrate a high level of formality, due to the usage of complex syntactic structures, especially the introduction of a high number of subordinate and non-finite clauses. Formality is as well obtained by the usage of the passive as well as by the lack of an explicit agent, which additionally guarantees the impersonality and objectivity of the text. The passive in such texts is particularly convenient, as it helps to thematize the data on which the texts concentrate. The neutrality of tone visible in the quoted texts constitutes the result of and, at the same time, emphasizes the informative purpose of the articles. Furthermore, the entries are characterized by expressive conciseness, the percentage of content words is higher, which contributes to greater lexical density and lower redundancy of the message. The expressions are shorter, yet at the same time clear and informative. Formality and neutrality are maintained also on the lexical level by the employment of sophisticated vocabulary, impersonal and non-emotive in tone.

The application of these syntactic and lexical means helps to obey the criteria of comprehensibility, conciseness and transparency within the articles.

To conclude, conventional entries are primarily informative and may be located closer to the formal end of the continuum. These articles exhibit a high level of discourse awareness and genre competence by the authors, who, despite the lack of any formal consequences in the case of an inappropriately written article, nevertheless strive to apply the generally accepted conventions related to academic discourse, so as to give Wikipedia a format and style typical of a traditional encyclopedia. Constructing and exploiting generic structures in a socially acceptable manner, the authors prove that they respect and maintain the norms and conventions of the genre. This brings them closer to the achievement of the communicative purpose associated with encyclopedic texts.

6.1.2. Unconventional articles

The second group of articles, comprising 638 entries, represents the collection of unconventional articles which manifest special and unusual discourse features. The differences concern textual, syntactic as well as lexical properties of discourse.

6.1.2.1. Textual features

There are a number of features which distinguish this group of Wiki-articles from traditional encyclopedic entries from the textual standpoint. The dominant properties of Wiki-entries in this respect encompass unconventional ways relating to textual organization, heterogeneity, exemplification, repetitiveness, informality and personalization.

6.1.2.1.1. Textual organization – integration versus fragmentation

Textual organization and appropriate paragraphing of the content help to assign a structural framework to the discourse, a framework which interrelates consecutive information, providing a perspective on the subsequent content (Brown and Yule 1983: 143), in a way that considerably enhances the degree of readability. The most frequent pattern of discourse organization to be found in encyclopedic texts is the general-specific pattern (Kister 1981, McCarthy 1991). According to this pattern, general information is presented first in the introductory sentence, with detailed data following in subsequent sentences or sections.

It is worth noting that tendencies towards structuring which are novel and originate from the distinctive environment in which the free encyclopedia is created can be observed. Wiki-texts exemplifying a departure from the conventions are characterized by a looser structure. The analysed material allows us to distinguish two dominant tendencies of content structuring, i.e. a tendency towards content integration and content fragmentation.

A. Content integration

Wiki-articles representing such an approach take the form of an unsegmented block of text. The authors do not isolate any specific paragraphs which are conceptually associated. Such articles resemble instead a stream of consciousness connected with the subject in question, the ideas being presented in a sequential fashion, but without the deployment of any typographical means of text-segmentation. Within the analysed group of articles, this approach was identified in 41.84% of the entries (267 articles).

The tendency can be seen in the extract quoted below. The author first presents a short introductory biography of Ulugh Beg and then moves on to a detailed discussion of chosen facts from the sultan's life. Thus, the reader is informed about Beg's family background and his rise to power. What follows is a lengthy and detailed exposition of Beg's scientific activity, especially his interest and achievements in astronomy. Next, the author returns to historic events and Ulugh's reign, before closing the discussion with further information associated with astronomy:

Ulugh Beg

Introduction – [Ulugh Beg was a Timurid ruler as well as an astronomer, mathematician and sultan. His commonly known name is not truly a personal name, but rather a moniker, which can be loosely translated as “Great Ruler” or “Patriarch Ruler” and was the Turkic equivalent of Timur’s Perso-Arabic title Amir-e Kabir. Ulugh Beg was also notable for his work in astronomy-related mathematics, such as trigonometry and spherical geometry. He was the grandson of the conqueror Timur and oldest son of Shah Rukh, bot of whom came from the Mongol Barlas tribe of Transoxiana. His mother was the Persian noble Goharshad. Like his grandfather Timur, he was also the ancestor of Babur, founder of Mughal Dynasty. Ulugh Beg was born in Sultaniyeh in Iran. As a child he wandered through a substantial chunk of the Middle East and India as his grandfather expanded his conquests in those areas. (...) In 1411 he became a sovereign of the whole Mavarannahr khanate. The teenaged ruler set out to turn the city into an intellectual center for the empire (...). His own particular interests concentrated on astronomy, and in 1428 he built an enormous observatory called the Gurkhani Zij. Lacking telescopes to work with, he increased his accuracy by increasing the length of his sextant (...). In 1437
Family back-ground	Ulugh Beg determined the length of the sidereal year as 365.2570370...d=365d6h10m8s. In his measurement within many years he used a 50m high gnomon. This value was improved by 28s88 years later in 1525 by Nicolaus Copernicus (...). Unfortunately Ulugh’s scientific prowess was not matched by his administrative skills. He lost some battles to rival kingdoms, and in 1448 massacred the people of Heart after defeating Mirza Ala-u-dowleh. Within two years he was beheaded by his own eldest son, while on his way to Mecca. In
Rise to power	honour of his achievements the Ulugh Beigh crater on the Moon was named after him by the German astronomer Johan Heinrich von Madler in his 1830 map of the Moon.
Scientific activity	(http://en.wikipedia.org/wiki/Ulugh_Beg, 15 th June 2007)
Administration of the country	
Death	
Respect of other astronomers	

The lack of segmentation in these cases does not lead to incoherence in the articles, but decreases the level of transparency of the content. Due to the considerable length of the articles (>800 words), a segmentation of the contents would definitely contribute to a greater degree of lucidity of the entries. Segmentation, at least in the form of indentation, would facilitate finding specific information in the text more quickly.

The flow-approach towards the composition of entries can be further exemplified on the basis of the article quoted below, where a lack of attention towards structuring is particularly noticeable. This instance presents a flow of information associated with the main subject, yet clearly in an unordered manner:

Ugaritic language

Historical information	The Ugaritic language is only known in the form of writing found in the lost city of Ugarit in Syria since its discovery by French archeologists in 1928. It has been extremely important for scholars of the Old Testament in clarifying
Usefulness for Hebrew texts	Hebrew texts and has revealed more of how Judaism used common phrases, literary idioms, and expressions employed by surrounding gentile cultures. Ugaritic was “the greatest literary discovery from antiquity since the deciphering of the Egyptian hieroglyphs and Mesopotamian cuneiform”. Literary texts discovered at Ugarit include the Legend of Keret, the Aqhat Epic, the
Contents of texts	Myth of Bal-Aliyan, and the Death of Baal, all revealing a Canaanite mythology. Ugaritic was a Semitic language written in cuneiform that was adapted for use as an alphabet. This Ugaritic alphabet, among the oldest that has been discovered, is different from all other cuneiform writings insofar as it is an
Definition/origin of the language/properties	alphabet rather than a syllabary. The so-called long alphabet has 30 different letters, while the short alphabet has 22. To the casual observer, it appears very similar in appearance to Akkadian or Assyrian writing. The Ugaritic language is attested in texts from the 14 th through the 12 th century BC. The city was destroyed in 1180/70 BC. Ugaritic was used by a Canaanite culture, and the use
Historical information	of the term ‘Canaanite’ to refer to the Ugaritic language is sometimes found. It is closely related to the Canaanite languages. (...). The study of Ugaritic is useful for biblical Hebrew scholars because the Ugaritic texts provide an unparalleled glimpse into the life and religious worldview of the ancient Israelites. The vocabulary is amazingly close to biblical Hebrew – many Ugaritic words
Usefulness for Hebrew texts – language/ religion	are letter-for-letter the same as biblical Hebrew. It is the religion of Ugarit, however, that is especially important to Old Testament scholarship, since Ugaritic is the ancient language of one of Israel’s closest neighbors, the city state of Ras Shamra, located in what is now Syria. (http://en.wikipedia.org/wiki/Ugaritic , 23 rd March 2007)

While in the previously quoted extracts paragraphs thematically associated can be singled out from the main body of the article, in this case, the text presents a flow of unsegmented pieces of information. An appropriate structuring of the message would enable the extraction of the most significant and relevant inter-related facts to give a clearer view of the subject. Thus, the sentences defining the Ugaritic language should be merged in the initial paragraphs, followed by the facts concerning its discovery, origin and relation to other forms of writing and concluded with the mention of its usefulness for the study of biblical Hebrew.

A conciseness within articles is typical of paper encyclopedias, where the entries are compressed and densely ordered due to space restrictions. Still, even in such publications, texts are segmented into paragraphs discussing specific areas of the subject in question. Segmentation is achieved by the use of different typographic devices, e.g. indentations. In this way, definitions become more coherent, since the data is arranged in a hierarchy of importance, relevance and reference.

Wikipedia, as noted earlier, in contrast to its paper antecedents, is not subject to spatial limitations and offers each writer an opportunity to segment the article into thematically associated paragraphs, which would give the entries a clearer and more readable form. As numerous Wiki-articles illustrate, the authors do not always take advantage of this facility.

B. Content fragmentation

The effect of content fragmentation, which was present in 27.27% of the unconventional Wikipedia articles (i.e. 174 entries), is a consequence of the content being split into short one or two sentence paragraphs. In this way, the articles are given the form of separate, isolated strokes, individual sentences, with the relationship between them not indicated. For instance:

Uniform prism

The term uniform prism can be used for a right prism with square sides since such prisms are in the set of uniform polyhedra.

A n-prism, made of regular polygons ends and rectangle sides approaches a cylindrical solid as n approaches infinity.

Right prisms with regular bases and equal edge lengths form one of the two infinite series of semi regular polyhedra, the other series being the antiprisms.

The dual of a uniform prism is a bipyramid.

A parallelepiped is a prism of which the base is a parallelogram, or equivalently a polyhedron with 6 faces which are all parallelograms.

A right rectangular prism is also called a cuboid or a rectangular box, while a right square prism is simply a square box, and may also be called a square cuboid.
(http://en.wikipedia.org/wiki/Uniform_prism, 16th November 2007)

In this case, each sentence carries an isolated piece of information associated with the concept discussed. These paragraphs are coherent to an extent, yet visibly lack in cohesion. Each sentence, constituting a paragraph of its own, refers to a different aspect of the issue in question. Moreover, the items of information are presented in a rather laconic, telegraphic manner, not providing a real explanation of the notion for a reader not knowledgeable in the subject.

The following article constitutes a further instance of the “split-approach” towards text-segmentation. In this case, all the sentences pertain to a description of a species, but appear in an unordered manner, with instances of repetition:

Uria

Uria is a genus of seabirds in the auk family. They are medium sized guillemots with mainly brown or black plumage in the breeding season. They breed on the coasts of the northern Atlantic and Pacific Oceans.

These birds breed in large colonies on coastal cliffs, their single egg being laid directly on a cliff ledge. They move south in winter to keep in ice-free waters.

They dive for food from the surface, swimming underwater. They mainly eat fish and crustaceans, also some molluscs, insects and plant material.

Adult birds are black or brown on the head, neck, back and wings with white underparts. The bill is long and pointed. They have a small rounded black tail. The lower face becomes white in winter.

The flight is strong and direct, and these species have fast wing beats due to the short wings.

Uria guillemots produces a variety of harsh cackling calls at the breeding colonies, but are silent at sea.

(<http://en.wikipedia.org/wiki/Uria>, 16th December 2007)

The reason for the split-structure of particular articles has to be attributed to the editorial freedom and collaborative character of the free encyclopedia, where each article is constantly subject to content additions and spontaneous insertions made by individual authors. The flow of paragraphs in such articles is abrupt, as successive authors insert new sentences into the structure in a random fashion, without paying attention to the effect their contribution produces on the overall structure and coherence of the article.

This practice is clearly reflected in the structure and content of the article below, where the introductory paragraph, comprising a chain of inter-related sentences, is followed by three successive one-sentence insertions:

Urs

Urs is the death anniversary of a Sufi saint in South Asia, usually held at the saint's dargah (shrine or tomb). South Asian Sufis being mainly Chishtiyya refer to their saints as lovers and God as beloved. They refer to their death as Wisaal (union with the beloved) and death anniversary as Urs (marriage).

The word 'urs' comes from the Arabic word for 'marriage'.

This Muslim religious fair which may include religious music, Qawwali.

The ceremonies are begun by the sajjada nashin (custodian of shrine).

(<http://en.wikipedia.org/wiki/Urs>, 12th December 2007)

The last three sentences clearly stand out from the entire structure, being spatially isolated from the introductory paragraph as well as from one another.

Poor paragraphing and the split structure of information is clearly evident in the following article, presenting the history of Umkirch, a German town. The account of the events is not presented in a coherent manner, with historical continuity interrupted by isolated sentences (e.g. Dachswangen), additions thematically diverging from the main account of the events (e.g. church). Moreover, the impression of incoherence and disharmony is enhanced by a lack of tense agreement in the description of the successive events:

Umkirch

Historically, it is first mentioned in 1087 in a document about the exchange of the monastery St. Ulrich in the Hexental between the Bishop of Basel, Vurkard, and the Monastery of Cluny in France. A witness, present at the deal his named “Humbert de Untkilcha. In 1270 Umkirch is given to the knight Dietrich Snewlin, later to several others, among them Martin Malterer, the count palatines of Tuebingen and Flora Coutness of Wrbna, until it was given to Grand Duchess Stephanie of Baden.

Insertion	Part of Umkirch is the village of Dachswangen: in 1924 it became part of the village.
Historical information continued, tense change	On April 20, 1945 Umkirch gets occupied by the French after World I Umkirch lost twenty one citizens, during World War II, thirty citizens. Twelve others are missing in war.
Another insertion, tense change	The first church in Umkirch was built by a Frankish magistrate, called Cenenarius Eliliant. He was also called ‘Hunde’. Apparently this name was given to the village: Hundechilche, meaning ‘Dog’s Church’. The church of Umkirch dates back into the second half of the 11. century. It is one of the oldest churches in Breisgau. The church was mentioned first in a letter, dated April 14, 1139 from Pope Innocent II, to Basel. (http://en.wikipedia.org/wiki/Umkirch , 19 th March 2007)

In this extract we may observe an interweaving of themes which relate to the town, the central topic, and themes which relate to different temporal frames and different historic events. Such structure visibly affects the theme-rheme sequence of the paragraph.

As can be observed in the examples, both types of textual organization may prove disadvantageous to the effectiveness of the articles, as in both the structure of discourse is not clearly marked. While the paragraphs which are tightly integrated may prove hard to digest for readers, those split into consecutive individual sentences may be perceived as insubstantial.

6.1.2.1.2. Heterogeneity

The “polyphony of the voices” of Wiki-authors is reflected at all the levels of the free-encyclopedia, with each of them producing interesting discourse properties. The magnitude of the authors and the editorial freedom are predominantly reflected in the overall heterogeneity of the discourse forms in particular articles. Heterogeneity goes beyond the conventional properties of the genre, which, as paper antecedents of Wikipedia illustrate, is characterized by standardization and homogenization of discourse.

The concepts of heterogeneity, hybridity, intertextuality or genre-mixing stem from Bakhtin’s (1986) claims for the intrinsically dialogical nature of every type of genre, be it spoken or written. As Bakhtin observes, any utterance or discourse draws on different yet multiple voices specific to a culture or society, and as such is not

a product of a single author, but a negotiation between various forms. Thus, the text may be monological in its formal aspect, but dialogical at the same time due to the implicit or explicit assimilation of other voices. Fairclough (1992) extends the dialogical conception of discourse, pointing to a considerable diversity of intertextual relations which texts may represent while exhibiting various degrees of heterogeneity, ranging from manifested and straightforward incorporation of different generic types in a text (e.g. the use of direct or indirect speech in texts), to a complex merger of various conventions, where intertextuality is not evident due to the tight stylistic and structural integration of the heterogeneous elements. The inter-generic relationships may sometimes involve the mixing of genres to the extent that a “radical deconstruction of genre boundaries” (*ibid.*: 90) occurs, and clear distinctions between different conventions become blurred, which might in consequence lead to the emergence of new types of genres (e.g. advertorial, infotainment).

The heterogeneity of the Wiki-discourse occurs at different levels and is expressed in the combination of features belonging to different genres and functional styles within one entry. To illustrate the different forms of intertextual relations of the Wiki-discourse, we propose the following categories of heterogeneity of the free encyclopedia:

A. Generic heterogeneity

Unconventional Wiki-articles reflect a play with genre patterns and different forms of stylization, when authors incorporate into the encyclopedic conventions forms associated with other genres. The presence of articles with complex and miscellaneous genre forms is conditioned by the fact that selected texts are often supposed to fulfil different functions.

Most frequently the entries incorporate the elements associated with promotional genres, i.e. advertisements, product marketing or brochures, elements of a tourist-guide. Such elements were identified in 13.47% of the articles (86 entries). In these cases, genre hybridization involves mixing of information with advertisement and persuasion. Such persuasion incorporated into Wiki-articles is expressed via different strategies.

Elements of persuasion can be observed in the articles in which writers do not simply focus on a description of the subject, but resort to a form of advertisement, trying to convince readers of the values associated with the concept described. In this case, the persuasive effect is achieved mainly by the use of evaluative adjectives, emphatic forms or an indirect appeal to the reader (e.g. by the use of *one*), as in the following examples:

Utwe

The park features excellent examples of healthy hard coral, as well as some of Micronesia's most pristine forests and extensive mangrove ecosystems.
(<http://en.wikipedia.org/wiki/Utwe>, 22nd December 2007)

Marble Rocks

These rocks are a gorgeous ones to watch in moonlit night. The journey in local ferry is an incredible experience. One can never afford to miss a trip to bedaghat.
(<http://en.wikipedia.org/wiki/MarbleRocks>, 27th July 2006)

In these examples, by introducing elements of evaluation, the authors not only inform, but also attempt to persuade the readers to accept their views concerning the subject.²

Such elements of persuasion can be observed predominantly in the articles closely resembling the form and content of a tourist-brochure. These articles do not focus solely on a description of geographical regions or cities, but also resort to distinguishing the special value of these regions, aiming to convince readers of their attractiveness, e.g.:

Udaipur Rajasthan

Bharatiya Lok Kala Mandal - An excellent museum of folk arts. Also hosts great puppet shows in its auditorium.

Saheliyon-ki-Bari- Maharana Sangram singh builds this in the mid 18th century. The ,garden of the maidens' brings to mind the lifestyle of the ladies of the court. The delightful gardens appear discreet and in impeccable taste. (...) The whole ambiances is flavored by the nostalgia of those beautiful bells enjoying themselves in a lavish aura.

Lake Palace - Originally known as the Jag Niwas, took three years to build and was inaugurated in 1746. Nothing but marble is used in the composition of this palace.

Shipgram - The crafts village, is a must see place for those who want to see the cultural heritage of India, its handicrafts, etc.

(http://en.wikipedia.org/wiki/Udaipur,_Rajasthan, 21st November 2007)

110

In the extract above, the author presents an enumeration of the beauty-spots of the city, places particularly worth seeing. Moreover, the author resorts to a strongly positive evaluation of these places. The description is essentially persuasive and is meant to promote the place, which is typical of advertising genres.

An interesting example of advertising was found in a biographic article on an American pastor. In this case, the persuasion is straightforward and interactive, as it takes the form of a direct inducement expressed through the use of imperative forms, explicitly addressing the reader:

Unabridged - Mike Moore

Dr. Mike Moore is seen and heard in cities throughout the nation on *The Word of God Is the Answer* weekly television broadcast, including Chicago, Little Rock, Miami and New York City. So tune in every Sunday with your pen, your pad, and your Bible to receive a life changing word from God.

Be on the lookout for Dr. Mike Moore's upcoming book "The Word Rich is not a Bad Word". Jesus came for us to have and experience the best this life has to offer, and Mike's book highlights biblical truths on how every Christian can experience a successful life.

([http://en.wikipedia.org/wiki/Michael_D._Moore_\(evangelist\)](http://en.wikipedia.org/wiki/Michael_D._Moore_(evangelist)), 9th March 2007)

The promotional character of the selected Wiki-articles confirms Bhatia's (1995) observation of an increase in the extent to which promotional strategies intrude into genres conventionally associated with a non-promotional, mainly informative pur-

² Subjectivity and evaluative remarks constitute a strong characteristic of Wiki-articles. This feature, manifested on the syntactic and lexical level of discourse, will be discussed in the subsequent sections of the analysis.

pose. The above-mentioned elements of persuasion point to a deeply rhetorical nature of the discourse of Wikipedia.

Apart from the elements of persuasion, Wiki-articles incorporate also other genre conventions. An interesting example of the Wiki-heterogeneity in this respect is the introduction of recipes into the entries, which were present in 19 articles (2.97%). For instance:

Ume Shu

1kg green ume

1kg rock sugar

1.8 liters of white liquor

Sterilise a 4 liter glass jar by filling it with boiling water, rinsing and drying carefully. Wash the ume, culling any fruit with bruises or broken skins. Dry the ume and remove the waxy bit in the stem end. Dry the fruit again (...).

(<http://en.wikipedia.org/wiki/Umeshu>, 31st March 2007)

The article on *Ume Shu* does not only provide a definition of Ume Shu (a drink), its origin and types, but also instructs the reader on how to prepare it. This extends the potential of Wikipedia, serving not only as the source of encyclopedic knowledge, but also as a source of practical information concerning specific subjects.

Ume Shu constitutes a very explicit example of Wiki-intertextuality, as it represents a mixture of different genre conventions. It not only provides an explanation of the concept together with the recipe, but also exploits elements of persuasion, thus incorporating three different textual forms. Persuasion in this instance consists in an indirect appeal to the reader by providing an evaluation of the merits of the drink, as visible in the following sentence:

Ume Shu

The taste and aroma of umeshu can appeal to even those people who normally dislike alcohol. (<http://en.wikipedia.org/wiki/Umeshu>, 31st March 2007)

A further interesting instance of heterogeneity is represented by the special stylization of articles, when authors give their articles the form of a story. Such stylization was identified in 8.46% of the articles (54 entries). In these cases, the shape of an article is dominated by an account of events, e.g.:

Ibn Hussain Ali Akbar

Ali al-Akbar ibn Husayn (Arabic: *علي بن الحسين*) (Fifth of Sha'aban, 44 AH - Tenth of Muharram, 61 AH) was the son of the third Shi'ah Imam, Husayn ibn Ali, and Umm Layla. He had a loud and beautiful voice, on the morning of the day of Ashura, Hussain ibn Ali asked Ali Akbar to recite the adhaan. Hussain ibn Ali, and many women in their tents, began to weep when Ali Akbar began reciting the Adhaan, knowing that it maybe the last time they are hearing Ali Akbar's adhaan.

Ali Akbar stood in front of Hussain ibn Ali after Zuhr prayers and said, „Father I request for permission to go and fight the enemies of Islam.” His father gave him permission and said, „May Allah be with you!

Hussain ibn Ali helped his son mount his horse. As Akbar began to ride towards the battlefield he heard footsteps behind him. He looked back and saw his father. He said, „Father, we have said good-bye. Why are you walking behind me?” Hussain ibn Ali replied, „My son, if you had

a son like yourself then you would have surely understood!” He killed many well-known warriors. No one dared to come close to him in a single combat.

(http://en.wikipedia.org/wiki/Ali_al-Akbar_ibn_Husayn, 26th September 2007)

Such articles have the form of a narrative, as they tell a story describing a particular event or character. These entries constitute an instance of heterogeneity different from the narrative elements described in Section C, since the narrative is not demarcated from the remaining body of the article and is not preceded by an informative section concerning the topic.

B. Stylistic heterogeneity

The collaborative nature of the free-encyclopedia is reflected not only in the differences between particular articles, but also in the internal diversification of individual entries. Here, collaboration is manifested in the mixture of different styles and registers within particular entries and reflects the individual styles of particular Wikipedians working on the editing of these articles.

The stylistic heterogeneity predominantly consists in the mixing of formal, academic and scientific register with colloquial and speech-like lexical and syntactic units. This form of heterogeneity is remarkably frequent in Wikipedia, as it was identified in 62.06% of the articles (396 entries). In the entries exemplifying such a form of hybridization, different parts characterized by heterogeneous linguistic elements may be identified, as can be seen in the following example:

UFO Defense: X-COM

A heavy weapons platform (HWP) is the general name for several military vehicles used in the game. HWPs are considered very useful: they have superior firepower to that of regular soldiers, and are less prone to damage from enemy fire. However, they cannot improve as the game progresses.

Researching a live Ethereal when your first encounter them, regardless of their rank, will let you to learn the skills of Psionics if you didn't already get it before. Soldiers give you info on the alien craft's mission. Medics give you autopsies or profiles of alien species. Navigators give you info on the Hyper Wave Decoder (if you don't already have it).

(http://en.wikipedia.org/wiki/UFO:_Enemy_Unknown, 16th March 2007)

The two paragraphs are taken from an article discussing a computer game and clearly differ in their stylistic shape, as they exhibit a juxtaposition of elements which have contradictory stylistic and semantic values. The first paragraph is written in standard English, using formal and impersonal means of expression. The second paragraph, in contrast with the former, introduces an evident change of style. The change is reflected in the oral tone of the presentation manifested in a higher degree of interactivity, consisting in the direct reference to the reader, the author's digressions as well as in the use of contractions. Thus, two main stylistic varieties are juxtaposed, the formal with the colloquial.

An analogous change in style within a mixture of formal and informal means of expression may be observed in the extract below:

Urban village

An urban village is an urban planning concept. Typically the aims are to create an environment which:

- People can live and work in
- Is self-sustainable
- Is environmentally friendly
- Encourages community development and bonding

In existing urban village developments, you can see buildings with shops on the ground floor and a mix of offices and apartments on upper floors. Hence, it would be possible to walk from your apartment to your office, which is just a few minutes down the street. Along the way you'd pass stores from which you can easily buy things, and not travel across multiple suburbs.

(http://en.wikipedia.org/wiki/Urban_village, 22nd March 2007)

This extract also exhibits structural heterogeneity, with its first part composed using enumeration and the second written in a descriptive manner.

The internal heterogeneity is even more clearly evident in the passage quoted below, due to an explicit transition between two different methods of description. The article opens with sentences characterizing Barbary Sheep, providing readers with the basic information concerning the origins, the habitat as well as the appearance of the animal. Yet, the description is interrupted by the introduction of an incoherent and nearly incomprehensible fragment. Though the insertion is thematically associated with the main subject of the article, it does not contribute to the informative value of the entry. On the contrary, it rather undermines its clarity and comprehensibility:

Udad

The Barbary Sheep is a species of Caprinae (goat-antelope) found in rocky mountains in North Africa. Six subspecies have been described. Although it is rare in its native North Africa, it was introduced in South Europe and other places.

Barbary Sheep stand 80 to 100cm tall at the shoulder and weigh 40 to 140kg. They are sandy -brown color, darkening with age, with a slightly lighter underbelly and a darker line on the back. Their horns have a triangular cross section. The horn curve outwards, backwards and inwards, and reach up to 50 cm. The horns are smooth but wrinkled at the base at some point if you touch the smoothness some where on the horn will be a little roughness. scientists say that the horn length can be about six inches or maybe more. but they are not sure but getting the accurate kind can be a little disastrous.

(<http://en.wikipedia.org/wiki/Udad>, 16th March 2007)

The second part of the article demonstrates not only greater informality of description, involving the use of colloquial vocabulary, hedge-like constructions, direct reference to the reader, but it is also characterized by a clear neglect of the basic rules concerning punctuation as well as the choice of vocabulary and syntactic constructions. The history page of this article shows that the section beginning with the words: *at some point...* (without a space between *baseat*) constitutes an addition made by a different user. This example shows that the authors collaborating in the creation of articles do not always have an awareness of the differences in style. This unawareness, as the extract illustrates, can, at times, verge on discourse ineptitude, where authors instead of contributing to an extension of the informative content of the articles, reduce the

level of lucidity. This particular extract constitutes also an obvious example of the violation of Grice's Maxim of Manner.

C. Marked heterogeneity

A marked heterogeneity in Wiki-articles assumes two major forms. These comprise the introduction of citations into the articles and the presence of narrative elements compositionally separated from the main structure of the articles. Intertextual relations in the form of citations and narratives are referred to as "manifest intertextuality" in Fairclough's (1992) terminology. In this form of intertextuality other texts are "explicitly present in the text under analysis; they are 'manifestly' marked or cued by features on the surface of the text" (*ibid.*: 104).

a) Citations

The high frequency of quotations, manifesting the intertextual nature of the free encyclopedia, constitutes a further distinctive feature, indicating a departure from a conventional expository and descriptive manner of content presentation. The frequent insertion of citations contributes to the polyphonic nature of Wiki-entries, as the passages quoted introduce different voices, opinions and, what is the most significant, different genre conventions into the texts of the entries.

Two main forms and origins of citations can be observed in the articles of the free encyclopedia, i.e. the quotations of spoken and written language.

The first major group of quotations concerns passages taken from spoken language. Within this group, frequent are quotations in the form of direct speech, which were present in 21.47% of the unconventional articles (137 entries). Such citations involve statements of an individual to whom the article is devoted or opinions concerning the individual uttered by other persons, e.g.:

Uh Huh Her

Harvey told *Mojo* magazine, "(...)Some of the songs are very gentle; with others I had a lot of fun in the words I used and the way I sang them." She explained to *Time Out* magazine how she "wanted to get back to the earthy, rootsy, more dirty side of things". She said, "I wanted this record to be simple, I wanted it to be ugly in some places, I wanted it to have a swagger to it." ([http://en.wikipedia.org/wiki/Uh_Huh_Her_\(album\)](http://en.wikipedia.org/wiki/Uh_Huh_Her_(album)), 23rd September 2007)

Udo Boelts

After Boelts had taken part in the Ironman Hawaii almost without preparation, his team chief during many years, Walter Godefroot, said that "Bolts is strong, he never breaks down". (http://en.wikipedia.org/wiki/Udo_B%C3%B6lts, 24th May 2007)

By giving a voice to others, the authors may incorporate into the entries different controversial statements and opinions, e.g.:

Umm Qasr

A British soldier was widely quoted in media reports as retorting, "There's not beer, no prostitutes and people are shooting at us. It's more like Portsmouth." (http://en.wikipedia.org/wiki/Umm_Qasr, 21st March 2007)

Udo Boelts

"Qual dich, du Schweine." (force yourself, you sod)
(http://en.wikipedia.org/wiki/Udo_B%C3%B6lts, 24th May 2007)

The quotations in biographical articles are frequently used with the aim of presenting an account of particular events in the past. In such contexts, the quotations are intertwined into the description, usually in the form of short dialogues inserted into the article, e.g.:

Udo Lattek

A dry spell in the domestic league in the 1974-75 season saw Lattek's tenure terminated prematurely and Bayern replaced him with Dettmar Cramer, who was also recommended to the club by Beckenbauer. "I told the president 'we need some changes'. 'That's right, you are sacked' he replied", Udo Lattek recalls this episode.

(http://en.wikipedia.org/wiki/Udo_Lattek, 19th November 2007)

This particular propensity towards providing reports of events is even more evident in the following extract from an article on Ernst Udet, a World War II pilot, where the quotations are interwoven into a description of specific episodes in his life:

Udet Ernst

Richthofen drove up one day as Udet was trying to pitch a tent in Flanders in the rain, and pointing out that Udet had 20 kills now, said:

"Then you would actually seem ripe for us. Would you like to?"

Of course Udet would.

(...)

When he excused himself for the WC the Germans fell over themselves trying to hide from him that they were watching to see if he would try to escape. On his return he said:

"I would never forgive myself for disappointing such hosts."

He did escape later.

(http://en.wikipedia.org/wiki/Ernst_Udet, 14th January 2007)

What is especially interesting in this example, is the style in which the author presents the account of the events, i.e. in a story-like fashion. Each quotation in this article is followed by a short sentence in the 3rd person, directly referring to the content of the preceding citation and acting as a commentary by the author, pushing the story forward and greatly dynamizing the description.

The inclusion of direct quotations helps the authors make the articles factual and more trustworthy, as in the following example, where the author of the article on the life of Ueshiba Morihei, a historical Japanese aikido master and practitioner, quotes passages from Morihei's diaries, which lends credence to the overall description:

Ueshiba Morihei

Ueshiba then walked to his garden and

"..I felt the universe suddenly quake, and that a golden spirit sprang up from the ground, veiled my body, and changed my body into a golden one. At the same time my body became light. I was able to understand the whispering of the birds, and was clearly aware of the mind of god, the creator of the universe."

His second experience occurred in 1940 when,

"Around 2am as I was performing misogi, I suddenly forgot all the martial techniques I had ever learned. The techniques of my teachers appeared completely new. Now they were vehicles for the cultivation of life, knowledge, and virtue, not devices to throw people with."

(http://en.wikipedia.org/wiki/Ueshiba_Morihei, 27th August 2007)

Similarly, the author of the entry below quotes a passage from a radio interview, the content of which is intended to confirm the author's statement concerning the person:

Uri Geller

In his telepathy demonstrations, Geller sometimes, but not always, reveals his answer slowly while asking whether he is on the right track. This approach is consistent with a stage magic technique known as cold reading, in which a magician tricks a subject into revealing information by suggesting that he already knows it. Geller's approach is apparent in an interview on the Gerry radio show on Feb 20, 2002:

Ryan: "are you getting the image that I'm sending to you? I'm concentrating very hard on it at the moment".

Geller: "It's very, very hard for me because, you know..."

Ryan: "just say what comes into your head, what's in your head?"

Geller: "Well the first thing that I drew was a ...it had a triangular shape at the top. Am I very wrong?"

Ryan: "I have sent you an image of the Pyramids. That's it! Are you really? You're not pulling my leg? No!"

Geller: "Gerry, I swear to you I drew a pyramid, and I also drew the stones in the pyramid, but I was not sure, so the first image that came into my mind was a triangle and then I drew the lines in it as the stones."

(http://en.wikipedia.org/wiki/Uri_Geller, 7th July 2007)

Apart from lending credence to a characterization, quotations in the extracts presented above considerably increase the vividness of the description, reaching beyond dry biographic accounts.

The citations introduced into the articles in the form of direct speech, as illustrated by the examples above, increase the level of informality, and transform the articles into a sequence of stylistically divergent utterances, resembling spoken mono- or dialogues. On the other hand, as can be observed in the aforementioned examples, quotations can make the articles more attractive to readers, as such citations as a rule do not appear in traditional encyclopedias.

In addition to the quotations taken from spoken language, Wikipedia articles contain a large number of citations from written texts. Such quotations appeared in 25.54% of the entries (163 articles). These are mainly excerpts from larger texts which constitute the main source of information for the content presented in the entries. For instance, in the following examples, the authors insert an excerpt from a United Nations military document and an extract from a medical document respectively, with the aim of authenticating the explication of the terms discussed and increasing the exactitude of description, thus making the articles more trustworthy:

UNEF

The first UN military force of its kind, its mission was to:

(...) enter Egyptian territory with the consent of the Egyptian Government, in order to help maintain quiet during and after the withdrawal of non-Egyptian forces and to secure compliance with the other terms established in the resolution...to cover an area extending roughly from the Suez Canal to the Armistice Demarcation Lines established in the Armistice Agreement between Egypt and Israel (...)

(http://en.wikipedia.org/wiki/United_Nations_Emergency_Force, 7th February 2007)

Umbilical cord

The Royal College of Obstetricians and Gynaecologists 2006 opinion states, “There is still insufficient evidence to recommend directed commercial cord blood collection and stem-cell storage in low-risk families.”

(http://en.wikipedia.org/wiki/Umbilical_cord, 6th February 2007)

The authors of articles focusing on the subjects associated with popular culture frequently resort to newspaper articles as the source of information, e.g.:

Uranium letter

The Times article stated that “According to senior European officials, in 1999 Rocco Martino provided French officials with genuine documents which revealed Iraq may have been planning to expand trade with Niger. This trade was assumed to be uranium, which is Niger’s main export.”

(http://en.wikipedia.org/wiki/Uranium_letter, 31st May 2007)

At times, as exemplified in 18 articles, the insertion of a quotation, however, increases the level of redundancy in the content, as the same information which is presented by the citation is successively repeated by either an authorial commentary or description, e.g.:

Umbar

Like the earlier New Haven in Enedwaith, and the later Pelargir on the Anduin, Umbar became a base from which Numenorean influence spread over Middle earth. It was at Umbar that the last king of Numenor, Ar-Pharazon the Golden, landed in 3261 S.A, to challenge Sauron:

The fleet came at last to that place that was called Umbar; where was the mighty haven of the Numenoreans that no hand had wrought. Empty and silent under a sickle moon was the land when the King of the Sea set foot upon the shore. For seven days he journeyed with banner and trumpet. Then he sent forth heralds, and he commanded Sauron to come before him and swear to him fealty. (“Akallabeth”, in The Silmarillion)

(<http://en.wikipedia.org/wiki/Umbar>, 29th April 2007)

Moreover, an accumulation of too many citations may impair the readability of the article, as in the quotation below, where the author incorporates the opinions of several scholars, all from different sources:

Unas

With his death, the Fifth dynasty came to an end, according to Manetho; he probably had no sons. Further, the Turin King List inserts a break at this point, which “gives us some food for thought”, writes Jaromir Malek, “because the criterion for such divisions in the Turin Canon invariably was the change of location of the capital and royal residence.” However, there are several clues of uninterrupted continuity between the Fifth and the sixth dynasties: Kagemni, the vizier of Unas’s successor Teti, began his career under Djedkare Isesi and Unas. Teti’s queen, Ipout is believed to have been the daughter of Unas, which shows Teti, Nicolas Grimal argues, “made no conscious break with the preceeding dynasty.” Jimmy Dunn adds that “a pink granite gateway in Unas’ mortuary temple bears the inscription of the names and titles of Teti, indicating that part of the temple was completed after Unas’s death.” The break between the two dynasties may have been more as an official act than in fact.

(<http://en.wikipedia.org/wiki/Unas>, 26th September 2007)

The extracts presented above constitute examples of citations separated from the main text by quotation marks or by different graphic means (e.g. italics, indentation).

The Wiki-discourse abounds also in the instances of indirect citations in the form of reported speech, which were found in nearly 30% of the articles (185 entries):

Uri Yehuda Ariel

Ariel said he believed that this would be correcting a historical injustice and that it is an opportunity for the Muslim world to prove that it is tolerant to other faiths.

(http://en.wikipedia.org/wiki/Uri_Ariel, 28th July 2007)

Utzon

In an article in Harvard Design Magazine in 2005, professor Bent Flyvbjerg argues that Utzon fell victim to a politically lowballed construction budget, which eventually resulted in a cost overrun on 1,400 percent. The overrun and the scandal it created kept Utzon from building more masterpieces. This, according to Flyvbjerg, is the real cost of the Sydney Opera House.

(<http://en.wikipedia.org/wiki/Utzon>, 21st March 2007)

Ubuntu

Ian Murdock, the founder of Debian, criticised Ubuntu in April 2005 for incompatibilities between its packages and those of Debian sarge, saying that Ubuntu had diverged too far to remain compatible.

(<http://en.wikipedia.org/wiki/Ubuntu>, 6th April 2007)

In these cases, the quoted messages are incorporated into the main body of the content by the use of reporting verbs (*say, argue, claim*, etc.). In these cases, citations are used to provide additional information concerning the subjects discussed, to show a different perspective, contradictory views or clashing opinions.

Taken together, the high frequency of occurrence of citations in the texts proves that they constitute a significant characteristic of the Wiki-discourse. The introduction of quotations is associated with an increase in the level of intertextuality and results in a change in the structure and generic value of the articles. The sources of the quotations represent a large and diversified collection of texts of different origins, genres and styles. As a result, the reader faces a hybrid, combining encyclopedic articles with interviews, official documents and literary and scientific texts. Still, intertextuality in this respect seems to be a positive phenomenon. The combination of the components of different origins appears to be valuable, interesting and contributes to the overall attractiveness of the articles, as it may trigger the readers' interest in and a greater involvement with the content.

b) Narrative elements

A further instance of a marked heterogeneity can be observed in the presence of narrative passages within the articles. Such elements, isolated from the main body of the entries as a separate textual structure, are introduced by metatextual cues.

The cues which signal the beginning of a narrative passage at the same time describe the kind of the narrative and its source. The most frequently used phrases include the following:

- “the story says”
- “there is a story that”
- “according to the narrative”
- “there is a legend”
- “the book tells about.”

The signals are placed in the sentence initial position, thus clearly marking the beginning of the narration. Narrative passages in this form appeared in 5.95% of the articles (38 entries out of 638) and may be exemplified in the following extracts:

Urith

Legend says Saint Urith was born at East Stowford in Swimbridge parish in the English county of Devon. She was a maiden devoted to the religious life. At the urging of an allegedly jealous, and perhaps pagan, stepmother, the local residents of Chittlehampton gruesomely beheaded the girl with a scythe. When she fell to the ground, a spring of water burst from the spot and flowers sprung forth (...).

(<http://en.wikipedia.org/wiki/Urith>, 15th January 2007)

Wilgefortis

According to the narrative, sometimes set in Portugal, a teen-aged noblewoman named Wilgefortis had been promised in marriage by her father to a pagan king. To thwart the unwanted wedding, she had taken a vow of virginity, and prayed that she would be made repulsive and in answer to her prayers she sprouted a beard, which ended the engagement. In rage, Wilgefortis's father had her crucified.

(<http://en.wikipedia.org/wiki/Wilgefortis>, 20th July 2007)

Ross Castle

There is a legend that O'Donoghue leaped or was sucked out of the window of the grand chamber at the top of the castle and disappeared into the waters of the lake along with his horse, his table and his library. It is said that O'Donoghue now lives in a great palace at the bottom of the lake where he keeps a close eye on everything that he sees.

(http://en.wikipedia.org/wiki/Ross_Castle, 6th May 2007)

Narrative passages of this kind are predominantly written in the past tense, using the third person singular form of verbs. Syntactically, narratives are presented in complex sentences, with a high level of coordination and the presence of conjunctions showing the sequential events of the story in question. As such, the narratives are separated from the main content only on the compositional level, without any accompanying structural changes, constituting a marked insertion into the main content of the entry.

The narrative passages are most frequently (49 articles, i.e. 7.68% of the entries) marked by metatextual cues in the form of a headline to introduce the narration. The use of headlines compositionally isolates the narrative passages from the main structure of the article. Headlines have the following forms:

- “Plot and/or ending details follow”
- “Story”
- “Spoiler warning.”

The thematic content of these articles includes films, books, computer games, anime series and cartoons.

A tendency which prevails in this type of narrative is the speech-like manner of presentation. Within the analysed material, 47.5% of all the narrative passages (28 articles) were written in a highly colloquial way, e.g.:

Tex

Easygoing, thoughtless, and direct, Tex at fifteen likes everyone and everything, especially his horse, Negro, and Johnny Collins's blue-eyed sister, Jamie. He thinks life with his seventeen-year-old brother, Mason, in their ramshackle house would be just about perfect if only Mace

would stop complaining about Pop. Pop hasn't been home in five months. Mace wants to get out of Oklahoma. Tex just seems to attract trouble and danger... Suddenly everything's falling apart.

([http://en.wikipedia.org/wiki/Tex_\(novel\)](http://en.wikipedia.org/wiki/Tex_(novel)), 9th January 2007)

Uzumaki

Kirie Goshima lives in a pretty normal town, until the day her classmate falls to his death from the top of a spiral staircase. Bizarre things start to occur; a girl's hair grows in an uncontrollable spiral pattern, a man kills himself by climbing into a running washing machine, a woman cuts out her inner ears and fingertips because they have spiral patterns, and a boy gets trapped beneath the tire of a car and twists in such a way as to become a spiral.

([http://en.wikipedia.org/wiki/Uzumaki_\(film\)](http://en.wikipedia.org/wiki/Uzumaki_(film)), 27th February 2007)

An even greater degree of informality can be observed in the narration taken from the article on *Relax*, an American movie. This extract has the form of a "spoken" reconstruction of the story in the film. For the sake of providing a fuller understanding of the tendencies reflected in Wiki-articles, we quote the whole narrative passage, as it best exemplifies the trends in the evolution of the Wiki-discourse, concerning not only narrative elements:

Relax

There is Murthy, the rich but frugal papa, and his two college-going kids, Siddhartha (Madhavan) and Shanti (Prema). Given papa's ways, the two kids feel a bit shackled, and yearn constantly for the good life. Matters come to a head when Siddharth wants to play the guitar at a college function, Shanti wants to take part in the fashion parade – both need money and, of course, papa isn't in the mood to indulge such extravagance.

Enter Raju (Abbas). A foundling brought up by Sumitra, who ekes out a living running an idle stall, Raju is a diehard Rajnikanth fan, working in a garage to help make both ends meet. Trouble begins when Raju loses his job, has a drink or three to ease the angst, and finds himself standing outside Siddharth-Shanti's home, gaping from outside at what he imagines is the rich life within.

At this point the family car rams into him. Raju is quickly carted inside the home by papa's servants – with a foreign delegation due any minute, the last thing he needs is embarrassment outside his gates. Raju is tucked up in bed. Not that there's anything wrong with him, but he figures he might as well play possum and enjoy as much of the 'rich' life as he can get.

The carefree lad makes friends with Sidharth and Shanti, listens to their vows, and promises to help them out. The treesome set of in papa's precious vintage car – and the movie from there on turns into a caper, revolving around the events of one frenetic day. Raju and Shanti become romantic, Siddharth pairs up with Suji, a baddie (Prakash Raj) comes into the equation, and a kidnapping goes wrong as Raju is mistaken for Siddharth.

How the foursome get out of the mess is what the rest of the movie is all about.

([http://en.wikipedia.org/wiki/Relax_\(film\)](http://en.wikipedia.org/wiki/Relax_(film)), 15th April 2007)

The orality of the passage quoted above is reflected both in its syntactic and lexical structures. Syntactically, informality is visible in such constructions as, e.g.: *There's Murthy, not that there's anything wrong with him, trouble begins*, etc. Lexically, informality is manifested in such lexemes as *gape, cart, ram into, lad, papa*, etc. as well as in the use of phrasal verbs: *get out of, be tucked up, eke out*, etc.

The passage, apart from its strongly colloquial form, exhibits the author's negligence of accuracy in the description, which can be observed in the variation in the spelling of one of the protagonist's name (*Siddharta*).

The colloquial character of such narratives is also shown in the extract below, which, additionally, reflects a strong authorial presence, due to the large number of digressive remarks incorporated into the content. Digressions, introduced in bracketed forms, or separated from the main body of the description by the use of dashes, provide additional information associated with the plot for readers unfamiliar with the story.

Moreover, the passage exhibits traces of ambiguity, introduced by the frequent use of personal pronouns, which do not have a clear reference, and as such may result in a reader's misinterpretation of the relations between particular characters in the story:

Uotani Arisa

Kyoko Honda, Tohru's mother, helped Arisa Uotani give up life as a member of an all-girl gang (Kyoko was an ex-gang member herself). Since then, she – affectionately called Uo-chan by her friends (see Japanese honorifics) – has been very protective of Tohru, and is grateful for Kyoko's influence. Her rough exterior is a misleading first impression, for close friends know her true generous spirit. In school, Kyo Sohma is her adversary (though they are very much alike in certain ways), and they bicker often. Due to her background, The Prince Yuki Fangirls find her scary and violent, and generally stay away from her.

(<http://en.wikipedia.org/wiki/Uotani>, 30th August 2007)

A characteristic feature of this type of narrative is the considerable dynamism of description. As the passages quoted below illustrate, this dynamism is achieved by a form of description where each sentence represents a successive event making up the story. Though the extracts differ in the tense forms of the verbs, they achieve a similar effect in producing a vibrant and lively description:

Ufo: Enemy Unknown

Story

In 1998, reports of UFP sightings began to increase dramatically. Tales of abduction and terrorism by the unknown aliens became widespread. The nations of the world came to perceive this as a threat and attempted to form their own forces to deal with this, such as Japan's Kiryu-Kai'. These forces failed miserably, the Kiryu-Kai not intercepting a single UFO in its five months of operation. On December 11, 1998, representatives from some of the most powerful nations in the world met in Geneva, Switzerland, to discuss the issue. From this meeting was born the Extraterrestrial Combat Unit – X-Com.

(http://en.wikipedia.org/wiki/UFO:_Enemy_Unknown, 26th March)

Ultrox

Xenoc, the heart of the universe, holds the valuable energy source known as Xeno-energy. (...) The protagonist, Ultrox, knows there is more to Xeno-energy than even the Zenterrans know. Just then, he notices a Kragan ship on a crash course. He saves its imprisoned occupant, Jek, from a squad of Klaw Troopers. Just before he was vapORIZED by another Kragan ship, the gangster named Seadrok accidentally dropped his space anchor on the ship, thus saving his life. After a little persuasion by dropping Ne'tal champion G'rog, Ultrox joins the races.

(http://en.wikipedia.org/wiki/Alien_Racers, 18th October 2007)

It is also worth mentioning those articles which abound in intertextual elements of different types. Within the analysed material, we have come across articles (5.95%)

which incorporate different types of intertextual elements and foreign generic conventions. One such example is the article on *Umbridge*, a character from the Harry Potter stories.

The article, apart from a definition and description of the character's features, tells the story of Umbridge, including the role she plays as well as the events in the specific volume of the book in which she features. All are presented in the form of a narrative inserted in the structure of the entry, e.g.:

Umbridge

Umbridge is subsequently installed at Hogwarts as the new Defence Against the Dark Arts professor by order of the Ministry of Magic. Her class consists only of defensive magical theory, due to Fudge's paranoid fear that Dumbledore intends to use his students as an army to bring down the Ministry. She is soon appointed the first (and only) ever "Hogwarts High Inquisitor", which gives her extraordinary powers over the students, teachers and curriculum.

(http://en.wikipedia.org/wiki/Dolores_Umbridge#Dolores_Umbridge, 30th December 2007)

Moreover, the author of the entry quotes passages from newspaper reviews concerning the character:

Umbridge

Novelist Stephen King, writing as a book reviewer for the July 11, 2003 Entertainment Weekly, noted the success of any novel is due to a great villain, with Umbridge as the "greatest make-believe villain to come along since Hannibal Lecter..."

(http://en.wikipedia.org/wiki/Dolores_Umbridge#Dolores_Umbridge, 30th December 2007)

As these instances of generic and stylistic heterogeneity prove, genre interference in the Wiki-discourse is multi-directional and results from the diversity of text-types and registers used in the composition of articles, and their contrastive juxtaposition in the entries.

Due to the introduction of generically foreign elements into the discourse, the above mentioned texts achieve the effect of originality. Yet, as a result, the conventions of the genre and its discoursal status become blurred. Such conscious contamination of the genre often leads to a modification of the pragmatic parameters of the articles, enabling the author to attain aims different from those conventionally associated with the genre of encyclopedia. As a consequence of the introduction of promotional elements into the entries, for instance, the communicative function of the articles ceases to be strictly informative, and evolves into persuasion. Genre heterogeneity, in this case, constitutes a direct consequence of the editorial freedom of Wikipedia, where each new entry can be created to fulfil a range of personal purposes linked to the individual contributors.

The heterogeneity of Wikipedia is closely associated with the concept of communicative competence. As some instances prove, the modification of genre norms may constitute an unconscious activity on the part of the author, arising due to a lack of competence and awareness of genre patterns.

6.1.2.1.3. Precision and exemplification

A feature characteristic of the free encyclopedia concerning its informativity is a clear propensity towards detail and minutiae as well as frequent exemplification of the issues in question. The attention to detail and extensive exemplification seem to be an attempt to enhance the readers' understanding of the concepts as well as to achieve a full and precise presentation of the issues.

However, the particularity and attention to detail observable in the free-encyclopedia can have both positive and negative consequences. Its positive aspect is associated with an increase in the comprehensiveness of the encyclopedia and its ability to satisfy the various inquiries of its users. As the extracts presented below illustrate, the details concern different types of information and different areas of knowledge. Most commonly, the attention to detail entails precision in the use of dates, names, numbers and statistics (characteristic of 13.63% of the articles), e.g.:

Upper Saxondale

The new building – designed by architect E. Purnell Hooley, better known as the inventor of Tarmac – was two stories high, cost 147,000 pounds and had accommodation for 452 patients (226 of each sex). The 130 acres surrounding the hospital cost 6,800 pounds. It was officially opened 24 July 1902 by Lady Elinor Denison. In 1913 extensions were made for 148 patients, which cost 29,833 pounds. In 1932, two further blocks were erected to accommodate 50 female patients each.

(http://en.wikipedia.org/wiki/Upper_saxondale, 23rd October 2007)

Upper gastrointestinal bleeding

Whiting studied a cohort of 325 patients and found the odds ratios for the strongest predictors were: black stool, 16.6 (95% confidence interval [CI], 7.7.-35.7); age < 50 years, 8.4 (95% CI, 3.2.-22.1); and blood urea nitrogen/creatinine ratio 30 or more, 10.0 (95% CI, 4.0-25.6). Seven (5%) of 151 with none of these factors had an upper GI tract bleed, versus 63 (93%) of 68 with 2 or 3 factors.

(http://en.wikipedia.org/wiki/Upper_gastrointestinal_bleeding, 15th July 2007)

Upfield Railway Line

Around mid-2004, peak services were increased from 3-carriage to 6-carriage trains. Previously 3-carriage trains running at 20-minute intervals could see the trains packed full of passengers. Frequencies vary according to the time of day and the day of week:

Daytime 7 days every 20 minutes

Evening 7 days every 30 minutes

Sunday early morning every 40 minutes

(http://en.wikipedia.org/wiki/Upfield_railway_line,_Melbourne, 22nd September 2007)

The precision of the articles in the free encyclopedia is as well enhanced by the existence of machine-generated articles. Namely, Wikipedia offers a large number of articles concerning geographical places, mainly cities and towns, which are generated by a software utility, a bot. Such software is specially programmed to extract the necessary information from a public census and automatically insert such data into a template. As a result, the content of such articles comprises mainly statistical information in the form of percentages, without any wider background information. Within the analysed material, the machine-generated entries encountered comprised

7.68% (49 articles). The bot-created articles present data predominantly associated with the United States, as in the following example:

Upper Marlboro

As of the census of 2000, there were 648 people, 292 households, and 165 families residing in the town. The population density was 610.2/km² (1,586.7/mi²). There were 309 housing units at an average density of 291.0/km² (756.6/mi²). The racial makeup of the town was 51.54% White, 45.06% African American, 0.46% Native American, 1.08% Asian, 0.15% from other races, and 1.70% from two or more races. Hispanic or Latino of any race were 1.54% of the population. (http://en.wikipedia.org/wiki/Upper_Marlboro,_Maryland, 16th October 2007)

but there is also a growing number of similar articles related to other regions of the world, e.g.:

Uttar Pirpur

As of 2001 India census, Uttar Pirpur had a population of 4789. Males constitute 52% of the population and females 48%. Uttar Pirpur has an average literacy rate of 63%, higher than the national average of 59.9%: male literacy is 69%, and female literacy is 58%. In Uttar Pirpur, 13% of the population is under 6 years of age. (http://en.wikipedia.org/wiki/Uttar_Pirpur, 21st October 2007)

Umm al-Fahm

According to CBS, in 2000, in the city there were 5,843 salaried workers and 1,089 are self-employed. The mean monthly wage in 2000 for a salaried worker in the city is ILS 2,855, a real change of 3.4% over the course of 2000. Salaried males have a mean monthly wage of ILS 3,192 (a real change of 4.6%) versus ILS 1,466 for females (a real change of -12.6%). The mean income for the self-employed is 4,885. There are 488 people who receive unemployment benefits and 4,949 people who receive an income guarantee. (http://en.wikipedia.org/wiki/Umm_al-Fahm, 26th October 2007)

The negative aspect of such fastidiousness is associated with the debatable value of such information. At times, the precision which the free encyclopedia seeks to attain comes down to the level of triviality. For the purpose of extending the level of informative content, Wikipedia authors frequently resort to the incorporation into the articles of trivial facts or details concerning the personal lives of the subjects, as in the following examples:

Uematsu

Uematsu currently lives in Japan with his wife, Reiko and his dog, Pao. (http://en.wikipedia.org/wiki/Nobuo_Uematsu, 9th November 2007)

Ulrich Salchow

Ulrich Salchow was married to the dentist Dr. Anne Elisabeth Salchow. (http://en.wikipedia.org/wiki/Ulrich_Salchow, 21st September 2007)

Upper Rissington

Upper Rissington's infrastructure is unusual in that its water and sewage are not connected directly to local supplies, but instead are operated privately by Reland, the village's developer. Residents pay Reland directly for these services instead of being under contract with a conventional supplier.

There is a small shop operated by the MidCounties Co-operative Society in the village. (http://en.wikipedia.org/wiki/Upper_Rissington, 15th November 2007)

Whether such facts should be included in the content of an encyclopedia constitutes an interesting point of debate, associated with the borderline between the so-called encyclopedic and non-encyclopedic information.

Moreover, apart from the above described particularity, Wikipedia, in comparison with a conventional encyclopedia, exhibits a marked tendency towards exemplification of the presented concepts. Exemplification, as Hyland (2007: 270) puts it “is a communication process through which meaning is clarified or supported by a second unit which illustrates the first by citing an example.” As the scholar observes, exemplification constitutes a vital aspect of exposition in academic writing, forming an instrument by which authors make their arguments more persuasive and accessible.

The rate of occurrence in Wikipedia (16.30%, i.e. 104 articles) proves that exemplification is equally essential for Wiki-authors. Wiki-authors frequently elaborate on the information they provide, restate the data and supply examples relating to their statements. By illustrating particular notions with examples, the authors manifest their awareness of their audience, that is readers without deeper knowledge of the subject, who turn to an encyclopedia for the purpose of acquiring information related to a specific topic. In this way, an increased degree of exemplification reflects the essentially dialogic character of the encyclopedia. Exemplification, as is evident in the extracts provided below, aims to facilitate the readers’ comprehension of the problem in question, and as such contributes to the effect of a “reader-friendly” text. For instance:

Cross-collateralization

This is the ratio of the amount borrowed against a property in relation to its value. For instance, a house that is currently valued at \$500,000 with \$200,000 debt has an LTV of 40%. This means the owner currently has borrowed 40% of the value of the property.

(<http://en.wikipedia.org/wiki/Cross-collateralization>, 24th August 2007)

Omission Bias

Spranca, Minsk and Baron extended the Omission Bias to judgments of morality of choices. In one scenario, John, a tennis player, would be facing a tough opponent the next day in a decisive match. John knows his opponent is allergic to a food substance. Subjects were presented with two conditions: John recommends the food containing the allergen to hurt his opponent’s performance, or the opponent himself orders the allergenic food, and John says nothing. A majority of people judged that John’s action of recommending the allergenic food as being more immoral than John’s inaction of not informing the opponent of the allergenic substance.

(http://en.wikipedia.org/wiki/Omission_bias, 19th September 2007)

Similarly, to clarify the meaning for their readers, the authors resort to the use of colloquial expressions associated with the particular concept, as in the following example, where the author rephrases the definition of the term using a common saying:

Train wreck

The term ‘train wreck’ is also used metaphorically to describe something distasteful or disastrous, yet inevitable, or something distasteful yet compelling in some form (‘you don’t want to stare, but you just can’t look away’).

(http://en.wikipedia.org/wiki/Train_wreck, 19th August 2007)

The examples presented above show that the authors try to shape meanings in such a way as to relate it to the readers' experience and their processing needs. Moreover, examples comprise references to more easily accessible concrete phenomena, usually associated with real life, thus facilitating the comprehension of the issues in question.

The authors resort to exemplification and supply additional information by elaborating on the issues not only for the purpose of explicating concepts, but also in order to provide an illustrative description of the issues which are the subject of an article. For instance:

UEFA

There was controversy when Liverpool won the competition in 2004/2005, but finished outside the top four in the Premiership. The Football Association ruled that Everton F.C. (who finished fourth) would get the final spot. However, UEFA came to an agreement that both Merseyside rivals would be allowed to enter the competition with Liverpool starting from the first qualifying round and Everton starting from the third qualifying round.

(http://en.wikipedia.org/wiki/UEFA_Champions_League, 16th December 2007)

Ueshiba Morihei

In his later years, he was regarded as very kind and gentle as a rule, but there are also stories of terrifying scoldings delivered to his students. For instance, he once thoroughly chastised students for practicing *jo* strikes on trees without first covering them in protective padding. Another time, as students sneaked back into the *dojo* after a night of drinking and brawling, he smashed the first one through the door over the head with a bokken, and proceeded to scold them.

(http://en.wikipedia.org/wiki/Ueshiba_Morihei, 27th August 2007)

The introduction of examples into Wiki-articles fulfils a primary role of providing a fuller perspective on the issues in question and facilitating the comprehension of the problems discussed. Examples enhance the educational value of the work, bringing the discussed issues closer to the readers.

6.1.2.1.4. Repetitiveness

While paper encyclopedias, due to space restrictions, are constrained in the scope of the information presented in the content, Wikipedia exhibits a propensity towards repetitiveness reflected in a tendency towards the repetition of the information already given, which could be seen in 24.60% of the articles (157 entries). This constitutes a consequence of an author's aspirations to maximum precision which, however, results in the extension of the articles and in the violation of Grice's Maxim of Quantity, as in the following instance:

Upgrade

The word is also used by audiophiles to describe the replacement of a product with a better quality product with the aim of bringing enhancements to sound quality.

The noun upgrade is used in audiophile circles to describe the replacement of a system component or components, for example a low quality or low powered electronic amplifier, with a better quality or more powerful amplifier from the same or different manufacturer's product range ostensibly to improve the quality of reproduced music from the hi-fi system.

(<http://en.wikipedia.org/wiki/Upgrade>, 15th October 2007)

Repetitiveness involves also an unchanged repetition of the same content, in different paragraphs of the articles. In the first instance below, the two paragraphs contain the same information expressed with similar wording, and with only a slight modification in the arrangement of the content:

Up for Grabs

Set in the booming international art market from 1990, which was fuelled by the dot com boom, it involves scenes of an alternated sexual nature. Not seen critically as one of Williamson strongest plays it is a play of bad manners, an analysis of how wealth and power can corrupt the arts. (...)

The boom was fuelled by the Dot com boom years, and Williamson also address six of the seven deadly sins, to address the ‘anything goes’ personal and sexual excesses of the time. A play of bad manners, an analysis of how wealth and power can corrupt the arts.

([http://en.wikipedia.org/wiki/Up_for_Grabs_\(play\)](http://en.wikipedia.org/wiki/Up_for_Grabs_(play)), 2nd September 2007)

Thus, the information appears in the entry first in the introductory paragraph, and is later repeated in the main body of the article.

A similar case may be observed in the instance below, where the information is repeated with the addition of a few words in the second paragraph:

University of Perpetual Help System

The University of Perpetual Help System DALTA has since subscribed to the institutional philosophy that national development and transformation is predicated upon the quality of education of its people. It is committed to the ideals of teaching, community service, and research, with “Character Building” as its guiding principle.

Then below:

The UNIVERSITY OF PERPETUAL HELP SYSTEM DALTA believes and invokes Divine Guidance in the betterment of the quality of life through national development and transformation, which are predicated upon the quality of education of its people. Towards this end, the institution is committed to the ideals of teaching, community service and research, with Character building is Nation building” as its guiding principle.

(http://en.wikipedia.org/wiki/University_of_Perpetual_Help_System_DALTA, 31st October 2007)

Such repetitiveness exemplifies the consequence of the collaborative nature of the free encyclopedia, with different users editing articles, and adding content without paying attention to the whole structure of the article or to the content already in existence.

6.1.2.1.5. Informality

Wiki-articles exhibit a marked and pronounced tendency towards orality, and a colloquial style of description, which has already been pointed out. Informality is reflected in the tendency towards dialogification in the message and an inclination towards expressivism, thus giving the articles the form of “written speech”. From among 638 articles, nearly half (278 articles, i.e. 43.57%) were written in a colloquial style. This can be observed in the article quoted below, which is written in a manner resembling a spoken reconstruction of events:

Triad

The Triad in its basic form consists of three members. The group is therefore independent of its members – if one leaves, there is still two left and the group will stay alive. The basic meaning of the word Triad is thus that it's a group that will survive its members. To the members of a Triad, the group is more important than their own personal careers.

Late 1986 Arrow and Fred of CoD were swapping intensively with guys all over the world. Typically they would receive 4-8 parcels a day with new stuff. Arrow got in touch with Ixion in 3001 and soon got to know they had a somehow stable access to originals. CoD and 3001 began cooperating for some time, as CoD had lots of contacts abroad, and 3001 had fast, short, high quality cracks. 3001 also knew Mr Z and his friend "the old man" RND. Now these three gangs (3001, CoD nad Mr Z + RND) with three distinct functions (Originals, Spreading and Cracking) saw a possibility of forming an independent Triad. As a result Triad was founded 1986-07-28 at 21.30 by Arrow, Fred, Skydive, Ixion, Lucifer, RND and at last, but no way least: Mr Z. On 1986-08-08 the heavy-metal lover Arrow and Fred spread a simple intro with a Swedish flag on top of the screen announcing the event.

When Ixion, who had been functioning as the main organizer of Triad, quit the scene to concentrate on his university studies in 1988 the golden triangle was handed over to 801DC. 801DC put down a lot of work to keep Triad alive during some shaky months before Jerry was taken into to group and the leadership was passed to him, with the blessing of the old founders. In 1991 801DC died in an airplane accident, missed by all his friends in Triad.

(http://en.wikipedia.org/wiki/Triad_C64, 3rd June 2007)

The informal character of the above quoted extract is manifested on a lexical stratum with the presence of colloquial words, e.g. *swap*, *shaky*, *guy*, *stuff*, *get in touch with*, *get to know*; idiomatic expressions, e.g. *with the blessing of*, *quit the scene*, *keep Triad alive*, hedges, e.g. *some*, *somehow* as well as contractions: *intro*, *it's*. Syntactically, informality is reflected in the use of complex, hypotactic and paratactic constructions, with a high degree of parenthetical clauses. It is manifested as well in the author's attention to detail, observable in the insertion of exact dates, names and a detailed description of concrete activities within the group.

Similar properties are noted in the lexical and syntactic structures which characterize the texts below:

Ultrix

Shannon and Stettner worked on low-level CPU and device driver support initially on UNIX/32V but quickly moved to concentrate on working with the University of California, Berkeley's 4BSD. Berkeley's Bill Joy came to New Hampshire to work with Shannon and Stettner to wrap up a new BSD release, incorporating the UEG CPU support and drivers, and to do some last minute development and testing on other configurations available at DEC's facilities. As an aside, the three brought up a final test version on the main VAX used by the VMS development group. No comments were heard from the VMS developers whose terminals greeted them the next morning with a Unix login prompt...

(<http://en.wikipedia.org/wiki/Ultrix>, 2nd April 2007)

Uriah Grant

In 1986 he defeated Matthew Saad Muhammad but was never really able to hit a groove. He looked headed towards a journeyman status, and dropped a series of fights, including a decision to Bobby Czyz in 1990. In 1991 he lost to Frank Tate, and moved up to cruiserweight. In 1993 he landed a shot at IBF Cruiserweight Title holder Alfred Cole and lost a decision. In 1995 Grant landed a rematch with Cole and lost another decision. Grant had a resurgence in 2000 when he shockingly TKO'd legendary but aging Thomas Hearns in the 2nd round.

(http://en.wikipedia.org/wiki/Uriah_Grant, 3rd May 2007)

In both extracts, the oral nature of the message is reflected in the accumulation of the lexemes referring to concrete, physical activities, names of objects, their properties, activities and states as well as in the incorporation of grammatical lexemes used to specify personal, spatial and temporal relations. As Chafe and Danielewicz (1986) note, place and time adverbials constitute signs of an author's involvement. Similarly, Biber (1986) interprets their presence as markers of a situated, as opposed to abstract, textual content. The colloquial character of the above quoted extracts is also reflected in the presence of discourse particles (*now, anyway*) and emphatics (*a lot, no way, really*), which mark involvement with the topic and are characteristic mainly of informal discourse and conversational genres (cf. Chafe 1982, Biber 1988).

Moreover, such texts are also characterized by a considerable degree of content personalization. As can be seen in the extract below, the cohesion of such articles is restricted due to a multi-thematic, digressive flow of discourse:

Upconverting

A video scaler is an algorithmic processing device for converting video signals between one arbitrary resolution/aspect-ratio and another resolution/aspect-ratio. In their most common application they are "upscaling" or "upconverting", taking a low resolution (Standard Definition) video and increasing the resolution to a high resolution (High Definition) video. This does not necessarily mean that the picture becomes clearer/more detailed - as scalers in their simplest form only increase the sample points for the original signal resulting in more data points for the original given information. Better scaling devices include other signal conditioning to maintain the original signal details when increasing the resolution. This also does not mean that placing a video scaler before a limited-capability display device will remove the limitations of that display device (for instance, you can't make a 720p display take a 1080p signal and expect to see all 1920x1080 pixels on the 1280x720 display surface).
(<http://en.wikipedia.org/wiki/Upconverting>, 17th July 2007)

The interactiveness, involvement and focus on personal attitudes in the above quoted extracts are reflected by features such as frequent pronouns, emphatics, inserted clauses and contractions. Owing to the presence of all these lexical and grammatical means, the extracts are characterized by greater empiricism, concreteness and factuality, dominant features of colloquial styles (cf. Biber *et al.* 1999).

On the one hand, this phenomenon may be treated as a violation of stylistic norms. Yet, on the other, such a high level of informality brings the author closer to the reader – the texts acquire the form of a spoken message directed at the reader. Informality and an emotional style strengthen the interactivity of the discourse, and constitute a sign of a hidden dialogue between the author and the reader. Such colloquial means also make the expressions more precise and vivid, carrying an emotional content. Such informality of discourse may also signify a wish to highlight its modernity and to purposefully differentiate the free encyclopedia from its traditional, conventional counterparts.³

³ The tendency towards informality is manifested as well through a deployment of isolated colloquial syntactic and lexical items present in the articles, which will be discussed in the sections below.

6.1.2.1.6. Personalization

The informal character of the unconventional Wiki-entries can be also observed in an authorial presence and personalization of discourse. The most prominent indication of authorial presence on the textual level is the accumulation of digressions.

Digressions are understood as intrusions in the flow of argumentation (Duszak 1997: 326), elements introducing content not directly associated with the main subject of a text. However, as Duszak (*ibid.*) emphasizes, digressions do not always contain content which departs from the main topic. They might comprise thematic inserts, with a focus on supplementation or explication of the main argument.

In the Wiki-discourse, digressions have the function of elaboration encompassing a restatement or clarification of the previously discussed material as well as amplification of the information. Digressions are introduced by signalling cues or in the form of parenthetical, bracketed segments. As such, they constitute a type of formal digressiveness (Duszak 1997: 327), with the digressions formally marked by text characteristics.

A. Signalling cues

The first group of digressions identified in the Wiki-discourse, present in 22.10% of the entries (141 articles), comprises authorial remarks introduced by signalling cues involving discourse markers and conjunctions. The following digressions are introduced in the article-final positions in order to comment on the preceding information. In all the cases, the digressions represent an authorial judgment or belief concerning the subject of the article:

Palestinian Municipal Elections

Over all, the local election showed Hamas relative strength and preparedness to deal with the block voting election system. On other side, it showed weakness and disorganization of Fatah and inability to understand the consequences of the voting system.

(http://en.wikipedia.org/wiki/Palestinian_municipal_elections,_2005, 27th January 2007)

Ubaldo Fillol

All in all, he is rated one of the best keepers of Football History.

(http://en.wikipedia.org/wiki/Ubaldo_Fillo, 28th June 2007)

Ammunition

However, worldwide statistics do not show unambiguously whether this number would be reduced by the prohibition of the sale of firearms and ammunition.

(http://en.wikipedia.org/wiki/Referendum_concerning_the_prohibition_of_the_sale_of_firearms_and_ammunition, 20th March 2007)

An authorial stance is even more clearly evident in the extracts quoted below. In the first two instances, the digressions are introduced with attitudinal adverbials, such as *amazingly* and *curiously*, respectively, thus reflecting the author's personal attitude towards the content, while in the third, the evaluation is reflected in the use of such adjectives as *useful* or *enjoyable*, which manifest the author's personal perspective concerning the subject:⁴

⁴ The use of lexemes expressing subjective opinions and content evaluation is discussed in Section 1.2.3.1.

Errors

Amazingly, they are able to easily understand all the other dialects of the Kalenjin, a feat which the others hardly accomplish.

(<http://en.wikipedia.org/wiki/Errors>, 19th November 2007)

Ubaldo Fillol

Curiously, Jan Jongbloed, the goalkeeper of the Netherlands, who played the 1978 World Cup final match against Argentina, didn't wear the number 1 either. He wore the jersey number 8, usually saved for midfield players.

(http://en.wikipedia.org/wiki/Ubaldo_Fillol, 28th June 2007)

Underwater diving

The ability to dive and swim underwater can be a useful emergency skill, and is an important part of watersport and navy safety training. More generally, entering water from a height is an enjoyable leisure activity, as is underwater swimming with or without breathing apparatus.

(http://en.wikipedia.org/wiki/Underwater_diving, 6th January 2007)

Both digressions, in contrast to those quoted above, constitute examples of authorial comments departing from the main thematic line of the articles.

B. Parentheses

Parentheses constitute a further specific signal of authorial presence in the structure of the text, more frequent than the previous type, as they occurred in 30.87% of the articles (197 entries). Bracketed remarks are introduced by Wikipedians for a variety of purposes, with their main aims being to supplement, restate or criticise the material presented in the main line of the argumentation.

Parenthetical remarks are the most frequent in the articles on science. In this context, they perform explanatory and elaborating functions, providing further clarifications of the processes in question and giving additional information or exemplification, e.g.:

Ueshiba Kisshomaru

After the war, and beginning in 1948, Kisshomaru oversaw the development of the Aikikai Honbu organization (and eventually the tearing down of the Kobukan Dojo in 1967 to construct the Aikikai headquarters).

(http://en.wikipedia.org/wiki/Ueshiba_Kisshomaru, 8th July 2007)

Decision problem

For example, it is possible for the graph of a function to be dividable in polynomial time (in which case running time is computed as a function of the pair (x,y)) when the function is not computable in polynomial time (in which case running time is computed as a function of x alone).

(http://en.wikipedia.org/wiki/Decision_problem, 21st April 2007)

Uptick rule

On the New York Stock Exchange a short sale may only be done on an uptick or a zero plus tick (which occurs when the price is the same price as the last trade, but higher than the previous different trade).

(http://en.wikipedia.org/wiki/Uptick_rule, 27th July 2007)

Here, clarification elaborates upon the meaning of the preceding statements to make them more transparent and accessible, and thus enhance the readers' understanding of the material.

A further function of bracketed remarks concerns the presentation of additional observations, subjective opinions, evaluations and authorial comments as well as remarks disclaiming the information presented in the main body of the article, e.g.:

Urban II

Roger I was to be free to appoint bishops, free to collect Church revenues and forward them to the papacy (always a lucrative middle position), and free to sit in judgment on ecclesiastical questions.

(http://en.wikipedia.org/wiki/Urban_II, 17th October 2007)

Umberto Bossi

Umberto Bossi is often described by some as a neo-fascist, racist (against immigrants and southern Italians, though his own wife is allegedly from Southern Italy, being actually a Lombard of Southern ancestry). Bossi agreed to return to an alliance with Berlusconi, which ultimately led to the (this time, easily predicted) 2001 electoral victory. (...)

(http://en.wikipedia.org/wiki/Umberto_Bossi, 19th February 2007)

Umanyar

It is also used as an insult by the sons of Feanor to Thingol (which is incorrect, since Thingol alone of the Sindar was reckoned under the Calaquendi).

(http://en.wikipedia.org/wiki/Umberto_Bossi, 6th January 2007)

132

However, additional information inserted in the parentheses does not always fulfil an explanatory function. Though intended as such, commentaries in brackets at times do not provide the readers with any sort of explanation, being imprecise and general in meaning:

Portals Athletic

Portals consolidated in Division 1 for two seasons until in 1986 they (along with the leagues top clubs with the required facilities) broke away to form the new Wessex League.

(http://en.wikipedia.org/wiki/Portals_Athletic_F.C., 2nd February 2007)

Peach Drop

(Georgia's nickname is "The Peach State," and Atlanta has a high proportion of streets and other geographical features named after peaches or peach parts as well as hosting the annual Peach Bowl).

(http://en.wikipedia.org/wiki/Peach_Drop, 5th September 2007)

It is clear in the last example that information is put unnecessarily within brackets, as it does not constitute an additional content, but belongs to the main body of the article.

Digressions, as exemplified above, may exert a positive or negative influence on the content. Some of the digressions presented above facilitate the understanding of specific problems by providing a further clarification of the presented notions. Other digressions, however, considerably extend the length of the articles, resulting in content proliferation and a greater level of repetitiveness.

6.1.2.2. Syntactic features

One of the most prominent features of encyclopedic discourse is its conciseness and the high concentration of information within an article.

Expressive conciseness may be achieved by the use of syntactic structures which increase the density of content and help to make the sentences more compact. These structures involve among others the use of nominalized forms, passive voice, a high frequency of content words and pre- rather than post-modifying structures.

While the majority of the afore-described conventional articles are created in line with the principle of expressive conciseness, the unconventional articles may be characterized by a departure from the rules in this respect. The syntactic structures which predominate in this group of articles are characterized by a low degree of nominalization, an infrequent use of passive constructions, compounding and an informality of construction.

6.1.2.2.1. Lack of nominalization

Nominalization represents a common syntactic structure specifically characterizing specialized forms of discourse. Traditional encyclopedias commonly resort to nominalized forms, since they allow a greater degree of precision and a greater density (Biber 1988: 104). Also, nominalized forms help to assign a greater degree of objectivity to an expression, which is especially significant with regard to encyclopedic articles.

Yet, in this respect, the Wiki-discourse departs from the convention, exhibiting a surprisingly low degree of nominalization. The authors do not strive to achieve conciseness and the economy of explication. The sentences are lengthy, with several coordinated elements, e.g.:

Udekkia

Japanese troops used Taiko drums to motivate troops, to help set a marching pace and to call out orders.

(<http://en.wikipedia.org/wiki/Udekkia>, 25th June 2007)

Ultrasonic flow meter

Ultrasonic open channel flow meters are widely used to measure flows in streams, rivers and where hydraulic head is an issue.

(http://en.wikipedia.org/wiki/Ultrasonic_flow_meter, 3rd October 2007)

Nominalization would seem especially appropriate in the cases of enumeration, which may be exemplified in the extract below:

Uday Hussain

Allegations include:

– Uday beat an army officer unconscious when the man refused to allow Uday to dance with his wife.

– Uday purchased or stole approximately 1.200 luxury vehicles.

(http://en.wikipedia.org/wiki/Uday_Hussein, 29th November 2007)

A low frequency of nominalized forms (present in 154 articles, that is 24.13%⁵) decreases the level of discourse formality. However, entries devoid of complex nominalizations acquire a more assimilable form, as a preference for verbal rather than nominal constructions makes them more readable and easier to comprehend for a wider circle of receivers.

6.1.2.2.2. Activisation

Conventional encyclopedic articles make considerable use of the passive. Passive forms are associated with a static, nominal style, and are considered the most important markers of a decontextualized and depersonalized style (cf. Biber 1988). On the textual stratum, passivisation increases the degree of coherence, especially when the passive is employed to thematize the process mentioned in the previous sentence.

Wiki-articles, on the contrary, exhibit a stronger preference towards activation of discourse, observed in 57.21% of the entries. Examples of this tendency may be seen in the extracts provided below. In these cases, rephrasals into passive forms would not only increase the level of formality, but would also move the focus towards the immediate subject of the article:

Ubv system

When people first started taking spectra of stars, they noticed that stars had very different hydrogen spectral lines strengths, and so they classified stars based on the strength of the hydrogen blamer series lines from A to Q. Later they found that some of the classes were actually duplicates and those classes were removed.

(http://en.wikipedia.org/wiki/Ubv_system, 6th April 2007)

University channel

Individuals may access the material directly from the University Channel website.

(http://en.wikipedia.org/wiki/University_Channel, 23rd April 2007)

The use of passive constructions seems especially fitting also in sentences where the authors do not explicitly state the agents, i.e. where the subject of the sentence has the form of an indefinite determiner. Such a situation can be seen in the examples below:

Udo Voigt

In 2005, Voigt compared the Bombing of Dresden in World War II to the Holocaust; some suggested this was a violation of Germany's laws on Holocaust denial.

(http://en.wikipedia.org/wiki/Udo_Voigt, 31st May 2007)

Ummah

Some modern Islamist use the term "Islamic Ummah" or "Muslim Ummah" to refer to all the people in the lands and countries where Muslim predominantly reside. They thus include non-Muslim minorities as member of the umma. When they talk of unifying the "Islamic Umma" they would include these non-Muslims, as citizen of the Islamic Umma. (...) Other Islamists accept the full equality of all citizens.

(<http://en.wikipedia.org/wiki/Ummah>, 14th January 2007)

⁵ The group comprises articles in which at least one or more instances of nominal forms were identified. This approach was applied in the assessment of the number of other syntactic structures as well, i.e. passive voice and embeddedness.

In these extracts the use of the passive would move the focus to the immediate subject of the article, as in these articles specifically it is not significant who expresses the opinions, but rather their content and reference is of importance. Such a transformation would result in a static and more abstract presentation of the information.

A low frequency of passive constructions and nominal forms reduces the lexical density of the constructions and decreases the level of detachment (cf. Biber 1988). This leads to an increased lexical prolixity and coordination, a further characteristic of the Wikipedic discourse.

6.1.2.2.3. Structure of sentences

Conventionally, as mentioned above, encyclopedic entries are characterized by reduced syntax (Sinclair 1988: 132-3, Görlach 2004: 175), which involves a rare usage of finite verbs, common words (conjunctions, articles) and frequent abbreviations. The style of unconventional articles, on the contrary, as already noted, is closer to spoken language, with its structure being less dense and less compact.

Wikipedia reflects a high proportion of sentences which exhibit a high number of interjections and authorial side notes. Such structures appear in 33.69% of the entries (215 articles), e.g.:

Ultramontanism

Some very extreme tendencies of a minority of adherents to Ultramontanism however, especially those attributing to the Roman Pontiff, even in his private opinions, of absolute infallibility even in matters beyond faith and morals, and impeccability, survived and were eagerly used by opponents of the Roman Catholic Church and papacy before the Second Vatican Council for use in their propaganda.

(<http://en.wikipedia.org/wiki/Ultramontanism>, 18th November 2007)

Unary input

Using a unary input is slower for any given number, not faster, the distinction is that a binary or a large base input is proportional to the base 2, or a larger base logarithm of the number while unary input is proportional to the number itself; so while the run-time and space requirement in unary looks better as functions of the input size, it is a worse function of the number that the input represents.

(http://en.wikipedia.org/wiki/Unary_numeral_system, 17th March 2007)

The sentences, with many additional comments, seem to constitute a way of achieving a precise description, as the authors try to integrate within a sentence as many details as possible, e.g.:

Ubastet

As a cat/lion goddess and protector of the lands when the fierce lion god Maahes became part of Egyptian mythology, she was identified, in the Lower Kingdom, as his mother.

As divine mother, and more specially as protectress, she became strongly associated with Wadjet, the patron goddess of Lower Egypt, eventually becoming Wadjet-Bast, paralleling the similar pair of patron and lioness protector for Upper Egypt.

This merging of identities of similar goddesses has led to considerable confusion, leading to some associating things such as the title Mistress of the Sistrum (more properly belonging to Hathor, who had become thought of as an aspect of Isis, as had Mut), and the idea of her as a lunar goddess (more properly an attribute of Mut).

(<http://en.wikipedia.org/wiki/Ubastet>, 15th September 2007)

A high degree of coordination, observable in the use of parenthetical clauses, reflects the oral nature of the unconventional Wiki-discourse, with its structure somewhat unplanned. Such constructions, however, may increase the effort the reader needs to put into extracting the intended meaning. The more compound the structure and the higher the number of digressions and analytic forms, the less transparent the meaning of the sentences becomes.

6.1.2.2.4. Epanaphora

Epanaphoric constructions involve the repetition of the subject or both the subject and predicate of the sentences, forming instances of excessive parallelism and recurrence, which can be seen in the following fragments:

Udinese Calcio

The club was relegated in 1961-62 and fell all the way to Serie C by 1964. The club remained in C for fourteen seasons before returning to B, and then reaching A in 1979. The club stayed in A until 1987 and then moved between the top two divisions for a time before re-establishing themselves in A from 1995.

(http://en.wikipedia.org/wiki/Udinese_Calcio, 11th November 2007)

Uccello

Paolo's first painting was a Saint Anthony between the saints Cosmas and Damianus, a commission for the hospital of Lelmo. Next he painted two figures in the convent of Annalena. Shortly afterwards he painted three frescoes with scenes from the life of Saint Francis above the left door of the Santa Trinita church. For the Santa Maria Maggiore church he painted a fresco of the Annunciation. In this fresco he painted a large building with columns in perspective. Back in Florence in 1446, he painted the Green Stations of the Cross, again for the cloister of the church Santa Maria Novella. Around 1447-1454 he painted Scenes of Monastic Life for the church San Miniato al Monte, Florence.

(<http://en.wikipedia.org/wiki/Uccello>, 11th April 2007)

Uros

They also hunt birds such as seagulls, ducks and flamingoes. and graze their cattle on the islets. They also run crafts stalls aimed at the numerous tourists who land on ten of the islands each year. They barter totora reeds on the mainland in Puno.

(<http://en.wikipedia.org/wiki/Uros>, 29th July 2007)

Such repetition, presumably, is to ensure the thematic continuity of the articles. Yet, in all the extracts presented above, repetition does not contribute to a greater degree of coherence within the message, rather increasing the redundancy of the discourse and decreasing the level of readability, as the reader may become distracted by the continuous repetition of the same forms.

6.1.2.2.5. Informality

The aforementioned informality of the Wiki-discourse is manifested also on its syntactic level, which can be seen in the syntactic structures typical of spoken language. As the tendency towards informality in the Wiki-discourse was described above, here we present only a few constructions exemplifying informality on the syntactic level. For instance:

Uvais Mohamed Imityaz

A classic rags to riches story, Imityaz was a Three Wheel Auto Driver from an impoverished background who contested local government election on an independent ticket. And at that stage he had no realistic chance of being elected.

(http://en.wikipedia.org/wiki/Uvais_Mohamed_Imityaz, 10th September 2007)

Battle of white mountain

Initially the revolt of the Protestants in Bohemia went well, and they broke out of their isolated political position by electing Frederic V, Elector Palatine as their king. But things changed when Maximilian I, Duke of Bavaria regrouped the forces of the Catholic League.

(http://en.wikipedia.org/wiki/Battle_of_white_mountain, 13th August 2007)

Ulrich Uli Stein

But there is a harder story: in the supercup of 1987, he gave Jurgen Wegmann, player of Bayern Munchen, a punch with the fist. After this scene, he was fired from Hamburger SV.

(http://en.wikipedia.org/wiki/Uli_Stein, 19th April 2007)

The colloquial character of the structures is reflected in elliptic constructions as well as in a high frequency of clauses beginning with conjunctions (e.g. *and*, *but*) (203 clauses in 127 entries – 19.90%).

6.1.2.3. Lexical features

With respect to the lexicon, unconventional articles exhibit a high level of subjectivity, interactivity and informality. It is worth noting as well the different uses of terminology in definitions. These features will be examined in detail in the following subsections.

6.1.2.3.1. Authorial presence – subjectivity and interactivity

One of the distinctive and remarkable properties of the Wiki-discourse, unlike the conventional encyclopedic language, is the emphasis on the authorial voice in the articles alongside various means enhancing interactivity in the entries. A sign of such involvement on the textual level assuming the form of digressions has already been mentioned in the previous section. However, authorial presence is reflected on the lexical stratum as well.⁶ The marking of personal authority in the articles involves self-mentioning and explicit remarks manifesting the authors' attitudes and judgments, while interactivity and the dialogic attitude are reflected in direct references to the readers.

A. Subjectivity

While in the majority of print encyclopedic texts authors remain invisible, detached from the message, in Wikipedia the authors frequently point to their presence in the articles. Although the process of writing is governed by an official principle of a neutral viewpoint, requiring the avoidance of bias and partiality, the authors do not refrain from presenting their personal attitudes as well as their intellectual and emo-

⁶ In the literature, such resources have been variously referred to, either as “appraisal” (Martin 1992), “evaluation” (Hunston/Thompson 2000) or as authorial “stance” (Hyland 1998).

tional evaluations. Despite the prescribed tendency towards impersonality, there are many purposefully used means specifically aimed at influencing the readers.

On the lexical level, subjectivism and the “I” of the author is commonly expressed in the evaluative vocabulary, which deprives the text of its neutrality and terseness.

In explicating complex scholarly issues, the authors frequently (28.84% of the entries) resort to remarks presenting their personal opinions concerning scientific facts or the phenomena described. While explaining concepts of science, the authors exhibit a propensity towards assigning a scale to the subjects according to their difficulty. To do this, they use lexemes such as *simple*, *easy*, *basic* or *clearly*. For instance:

Triangular matrix

A matrix equation in the form $Lx=b$ or $Ux=b$ is very easy to solve.

(http://en.wikipedia.org/wiki/Triangular_matrix, 8th April 2007)

Usonian

The word is clearly cognate with the Esperanto name for the United States, *Usono*, from whose root form *Uson-* one also gets the adjective *usona*.

(<http://en.wikipedia.org/wiki/Usonian>, 25th August 2007)

Specific Heat Capacity

Of course, from the above relationships, for solids one writes $C_m=C/m=cv/p$.

(http://en.wikipedia.org/wiki/Specific_heat_capacity, 7th November 2007)

These forms allow writers to evaluate the concepts defined and express their individual perception of the problems in question.

Such an approach can be observed in the example provided below, where the author evaluates the presented phenomena, placing them on a scale with respect to their likely occurrence and, at the same time, suggesting the reader's unfamiliarity with the concepts discussed:

Cecidomyiidae

Cecidomyiidae are also known for the strange phenomenon of paedogenesis in which the larval stage reproduces without maturing first. Even stranger in some species the daughter larvae produced within a mother larva consume the mother and in others the reproduction occurs in the egg or pupa.

(<http://en.wikipedia.org/wiki/Cecidomyiidae>, 31st March 2007)

Evaluative judgments, both positive and negative, are frequently presented in the thematic part of the sentences (113 instances), which gives the information a greater degree of acceptability (cf. Starzec 1999). In such a manner, evaluation achieves the form of an explicit opinion, an obvious observation, e.g.:

Uriankhai

It is a false notification that Tuvan and Mongolian language or even Turkic and Mongolian language are two different languages.

(<http://en.wikipedia.org/wiki/Uriankhai>, 23rd September 2007)

Specific Heat Capacity

It is not surprising that constant pressure heat capacities may be defined in terms of enthalpy.

(http://en.wikipedia.org/wiki/Specific_heat_capacity, 7th November 2007)

Urinary system

The kidneys receive their blood supply of 1.25L/min from the renal arteries which are fed by the Abdominal aorta. This is important because the kidneys' main role is to filter water soluble waste products from the blood.

(http://en.wikipedia.org/wiki/Urinary_system, 26th October 2007)

Negative evaluations in the analyzed articles are considerably strong and direct. The authors of the articles reject certain beliefs in a straightforward manner, explicitly negating particular views. For example:

Undecagon

The same shape is sometimes called an undecagon, but this is incorrect – the prefix should be the Greek 'hen', not the Latin 'un'.

(<http://en.wikipedia.org/wiki/Undecagon>, 12th February 2007)

Specific Heat Capacity

It has been assumed that atoms have no rotational or internal degrees of freedom. This is in fact untrue.

(http://en.wikipedia.org/wiki/Specific_heat_capacity, 7th November 2007)

Cecidomyiidae

Worldwide there are 3000+ species but since 1,100 are from well-studied North America it is clear this is an underestimate.

(<http://en.wikipedia.org/wiki/Cecidomyiidae>, 31st March 2007)

Evaluation of the content can also be seen in the emphasis authors place on certain parts of the message by highlighting specific fragments of the text. Frequent in this role are modal verbs, referring to necessity and obligation, such as: *we should consider, we should mention* (106 occurrences). They do not enable a freedom of interpretation, as the significance of specific facts is directly emphasized, e.g.:

Übermensch

It should be emphasized that the obstacles in becoming Übermensch are essentially internal, a matter of overcoming oneself.

(<http://en.wikipedia.org/wiki/Übermensch>, 29th November 2007)

Umag

We should not forget to mention the fishing trade that is of vital importance to the area and has been widely spread, too.

(<http://en.wikipedia.org/wiki/Umag>, 30th May 2007)

The emphasis put on these fragments reflects the author's hierarchization of the content. As such, the emphasis constitutes a means of evaluation, it highlights the fragments, which in the author's opinion are particularly significant to the process of understanding the problems in question.

The subjectivity of Wikipedia articles is visible not only in the authors' expressions of personal opinions concerning the difficulty or validity of scientific theories, but above all in the manifestation of individual views on popular subjects. It is mainly in the articles from the domain of popular culture that lexical evaluation is particularly strong and explicit. The most frequent mechanism of persuasion is the stimulation of a reader's attention through the evaluation of various phenomena, at times intertwined

with emotional colouring. The most frequent means of ascribing value to an object is the use of adjectives and adjectival phrases with emotive connotations, such as *interesting* (284 occurrences in the analysed group), *remarkable* (94), *beautiful* (87), *incredible* (62), with which authors manifest their personal attitude. Moreover, the authors resort to the use of lexemes with a particularly strong emotional charge, such as *breathhtaking* (15) or *gorgeous* (8), as exemplified in the extracts presented below:

Uzere

But these views or historical accounts, albeit oral, make interesting reading.
(<http://en.wikipedia.org/wiki/Uzere>, 4th April 2007)

Spanking

It is remarkable that English has a specific, common verb for spanking, not for punishment applied to other parts of the anatomy.
(<http://en.wikipedia.org/wiki/Spanking>, 25th November 2007)

Uzumaki

Among many bizarre features of the film is a heavy treatment with green colour filters.
([http://en.wikipedia.org/wiki/Uzumaki_\(film\)](http://en.wikipedia.org/wiki/Uzumaki_(film)), 27th February 2007)

Umtali

The view from the top of Christmas Pass down onto Mutare is breathhtaking.
(<http://en.wikipedia.org/wiki/Umtali>, 2nd August 2007)

Marble Rocks

These rocks are a gorgeous ones to watch in moonlit night.
(<http://en.wikipedia.org/wiki/MarbleRocks>, 27th July 2006)

Such adjectives indicate the reliability of the statements, in positive terms. Moreover, apart from manifesting such a conviction, they are also interactive in character (Biber 1988: 106). By means of the aforementioned lexemes the authors not only describe their feelings, but also significantly influence the readers' perception of the content.

The emotional tone of the text is also emphasized by reinforcement. In 98 entries, i.e. 15.38% of the articles, the authors chose to intensify the emotive load of the message by adding a number of terms with emotional or evaluative connotations as well as by the use of adverbial amplifiers (e.g. *really* – 138):

Uzeda

Producer Steve Albini recorded three of their records, and significantly contributed to the evolution of their sound: fiery, nervous, sludgy, and unpredictable, with sudden time changes and wild dynamics.
(<http://en.wikipedia.org/wiki/Uzeda>, 21st April 2007)

Zhu Tong

His depiction of a fatherly character of the young boy is heart warming and deeply moving, and one really begins to feel compassion for Zhu after Li Kui kills the young boy.
(http://en.wikipedia.org/wiki/Zhu_Tong, 15th May 2007)

Udaipur

The roads in this part of the country are really nice, paved in the middle, and fit for self-driving.
(http://en.wikipedia.org/wiki/Udaipur,_Rajasthan, 21st November 2007)

These lexical means express the author's evaluation, their emotional involvement and their personal standpoint on the issue.

In 14.89% of the articles (95 entries), the authors convey personal criticism in an uninhibited manner, overtly referring to themselves in the texts and openly presenting their personal beliefs. For instance:

Usui

In the absence of any direct knowledge of Dr. Usui's life the myths and legends were born. And some are quite amazing. I have a photocopied set of notes from another Reiki master in which Dr. Usui's healing miracles are somberly and reverently recorded.

(http://en.wikipedia.org/wiki/Mikao_Usui, 17th September 2007)

Ufology

Many ufologists are amateurs, and however well intentioned, are often unfamiliar with generally accepted research standards, thus often rendering their own research useless even to sympathetic mainstream experts.

(<http://en.wikipedia.org/wiki/Ufology>, 14th September 2007)

The accumulation of the above-mentioned evaluative strategies can be further observed in the following extract, where in one sentence, the author uses four means of assessment, adjectives and adverbs (*relatively simple, excellent, surprisingly, strongly*):

Bitter Sweet

The relatively simple plot – set in 19th century and early 20th century England and Austria-Hungary, and conceiving a young woman's elopement with her music teacher – is used as the hook for a series of excellent light-operatic numbers, many with surprisingly complex melodies strongly reminiscent of the best work of Gilbert and Sullivan.

(http://en.wikipedia.org/wiki/Bitter_Sweet, 10th February 2007)

Such instances of evaluation are clearly against the principle of neutrality, which is explicitly stated in the guidelines of the free encyclopedia (i.e. the Neutral Point of View Principle) as its basic tenet.

It seems that the authors of Wikipedia articles strive to achieve different aims, which entail not only the presentation of facts and the transfer of information, but in many cases involve also the manifestation of personal knowledge on a particular subject.

The presence of evaluative remarks manifests a departure from the compositional and linguistic conventions of the genre. The accumulation of emotive lexis induces a change in the communicative purpose of the text – the primary purpose of an encyclopedia, which is to inform about particular phenomena, is here supplemented by a personal interpretation of these phenomena.

B. Probability and conjecture – use of hedges

It has been observed (Biber 1988) that markers of subjectivity (e.g. emphatic means), frequently co-occur with hedges. Such co-occurrence is unequivocally confirmed by the Wiki-discourse properties.

Despite a large number of studies devoted to the phenomenon, hedging is not easily defined. Lakoff (1972) defined hedges in terms of truth values, by stating that they help to place concepts along a continuum between truth and falsity. As Chafe and Danielewicz (1986) and Biber (1988) argue, the main function of hedges is to

indicate probability and a lack of precision in the presentation of information, and as such serve as the means of expressing epistemic modality, i.e. modality which concerns “the opinion and attitude of the speaker” (Lyons 1977: 452), so judging the truth of a statement in terms of possibility or certainty (Hyland 1998: 44).

Hedging has been observed to enhance the interactional aspect of discourse, introducing a more reader-oriented perspective. As Hyland (1998: 246) argues, hedges are employed to acknowledge the reader’s right to refute the claims made by the author. This means that by modifying the claims and statements, by indicating uncertainty and tentativeness, an argument becomes more open to discussion, to acceptance or rejection on the part of the reader.

Hedges are commonly expressed through lexical verbs, modal verbs, epistemic adjectives and adverbs. Nearly all of these means are found in Wikipedia, where they mainly seem to encode authorial reservations as to the extent of personal commitment that the authors wish to devote to a statement. Worth noting are instances in which the authors resort to conjecture, speculation or inference drawn from their general knowledge or information obtained from other sources. Frequently, the authors avoid ascribing a complete truth value to the presented content. In this context hedges mainly work as a protective shield against the possibility of the data being considered mistaken.

a) Adverbs

142

The most common adverbs used as hedges present in Wikipedia are lexemes indicating tentative epistemic modality, i.e. those adverbs which express a degree of doubt or state the sense in which a speaker judges the statement as true or false (Quirk *et al.* 1985: 620). The most frequent are: *probably* (254), *maybe* (203), *perhaps* (123), *supposedly* (114), *reportedly* (71) as well as *presumably* (64) and *likely* (63), all expressing degrees of likelihood. For instance:

Weinstein

He was reportedly arrested for assault in the Netherlands and deported to the US. In 1972 he reportedly murdered one of the other inmates in the halfway house.
(http://en.wikipedia.org/wiki/Raymond_Weinstein, 6th June 2007)

Umkirch

The settlement exists since Roman times and was called Ecclesia in Undis (Church beneath the Waves – maybe because of the two small rivers floating passing though the village).
(<http://en.wikipedia.org/wiki/Umkirch>, 19th March 2007)

White Lotus

The ill-organized rebels managed to defeat the inadequate and inefficient Imperial forces, (...) perhaps contributing to the greater frequency of rebellions in the 19th century.
(http://en.wikipedia.org/wiki/White_Lotus, 7th January 2007)

The adverbs used in the above quoted extracts have a mitigating and downtoning effect and introduce the impression of speculation. These adverbs express a degree of tentativeness with regard to the statement, indicating doubt and a lack of complete certainty as to the applicability of the presented information.

b) Some

A lack of precision can be seen in the frequent use of “vague expressions”, such as *some* (472), which reduce the force of the statement, as in the following examples:

Uzbek Som

Some successor states had their national currencies before the change, some chose to continue using the pre-1993 Soviet/Russian ruble, and some chose to use both the pre-1993 and the new Russian ruble.

(http://en.wikipedia.org/wiki/Uzbek_som, 9th June 2007)

St. Baithen

Some writers assert that St. Baithen of Iona is the patron of Ennisboyne.

(http://en.wikipedia.org/wiki/Baith%C3%A9ne_mac_Br%C3%A9naind, 26th January 2007)

Decompression sickness

There are some reports indicating a higher risk of altitude DCS with increasing age.

There is some indication that recent joint or limb injuries may predispose individuals to developing “the bends”.

(http://en.wikipedia.org/wiki/Decompression_sickness, 28th November 2007)

The use of such lexemes suggests a low degree of informativity and directness. The authors of the examples presented above provide factual data concerning the subjects in question, yet their explications are devoid of precision, as they do not state explicitly the sources from which the information is obtained. In these extracts, *some* may indicate that the authors either consider the presentation of a source as redundant, or are not able to quote them precisely due to insufficient knowledge on the subject.

c) Modal auxiliaries

Among modal auxiliaries with a hedge-like effect, *may* is the most frequent in Wiki articles (266 occurrences). It is considered a prototypical hedge, indicating a 50-50 assessment of possibilities (Hyland 1998: 116). Similarly to the previous instances of hedges, it has a downtoning effect. In the following extracts, *may* is accompanied by the aforementioned modal adverbs, which additionally confirms their frequent use:

Corralito

In mid-2001 it was probably clear to bank owners and high-ranking officials that Argentina’s banking system was going to crash, and some in fact may have spurred this outcome by letting their highest deposit holders know this news.

(<http://en.wikipedia.org/wiki/Corralito>, 18th September 2007)

Upotipotpon

The name perhaps may have been derived from Aboriginal words *pootong pootong* meaning plenty of grass.

(<http://en.wikipedia.org/wiki/Upotipotpon>, 7th August 2007)

Ulver

Neo-classical composer and film scorer Craig Armstrong may have been an influence on Ulver as his use of electronics, trip-hop and hip-hop beats over strings and pianos is somewhat reminiscent of Ulver’s later works.

(<http://en.wikipedia.org/wiki/Ulver>, 20th November 2007)

Apart from *may*, a mitigating effect is achieved by the use of *could* (63 usages), as in the example below:

Uesugi Kagekatsu

The Sekigahara Campaign could be said to have begun, at least in part, with Kagekatsu, who was the first daimyo to plan revolt against Tokugawa Ieyasu.

(http://en.wikipedia.org/wiki/Uesugi_Kagekatsu, 5th June 2007)

Modal verbs, as with other hedging constructions, help the authors weaken the claim of absolute truth in the presented content. Also, they indicate that the writer is presenting a subjective interpretation of particular facts. These verbs as well give the readers more freedom and scope in their interpretation of the data.

d) Copulative verbs

Seem and *appear* represent another strategy used in hedging (Chafe 1985), their main function being to mark perception (Quirk *et al.* 1985: 1183), rather than a bald assertion of facts. These verbs also express the author's lack of complete commitment towards the presented content and reflect an impersonal uncertainty regarding a statement. Within the analysed material there were 86 occurrences of *appear* and 75 of *seem*, e.g.:

Urith

A book of her life, containing a record of her miracles, was at one time present in her shrine, and appears to be the basis of a rhyming poem in Latin now held by Trinity College, Cambridge.

(<http://en.wikipedia.org/wiki/Urith>, 15th January 2007)

Urban II

Roger I as a secular ruler seemed a safe proposition, as he was merely a vassal of his kinsman the Count of Apulia, himself a vassal of Rome, so as a well-tested military commander it seemed safe to give him these extraordinary powers, which were alter to come to terminal confrontations between Roger I's Hohenstaufen heirs.

(http://en.wikipedia.org/wiki/Urban_II, 17th October 2007)

The choice of these verbal forms introduces an effect of hesitancy and serves to distance the writer from the statement. The author mitigates the statement, at the same time allowing for the modification of the presented views and admitting that other solutions are possible.

An accumulation of the lexemes with a hedging effect is illustrated in the extracts presented below, where the authors do not provide the source of the information they quote, thus leading to speculation concerning the statistical data:

Uzbekisation

According to some sources, The Tajik population of Uzbekistan may be as high as the 42% of the total population of the country.

(<http://en.wikipedia.org/wiki/Uzbekisation>, 14th January 2007)

Umbrella Murder

However, it seems that the sugar coating of the bullet protecting the ricin content was damaged during the shot or before, and thus, only a tiny portion of the poison got into his blood, causing fever only. (...). However, experts declared that the probable gun that shot the bullet was probably very sophisticated, another reason to believe in State action.

(http://en.wikipedia.org/wiki/Umbrella_Murder, 3rd September 2007)

In the examples presented above, the hedges reflect the authors' judgments concerning the reliability of the presented information. They carry important information about the authors' stance, expressing their uncertainty about the content provided. Such an approach seems to be at variance with the principle of exactitude guiding the production of an encyclopedic work. A high frequency of lexemes producing an effect of uncertainty constitutes a feature going beyond the discursual characteristics of a traditional encyclopedia, where the factual information as well as the sources of data should be presented in detail, leaving the reader in no doubt concerning the accuracy of information.

Hedging constructions have been observed to constitute characteristic properties of overtly argumentative and persuasive genres, such as research articles (Hyland 1998, 2001), where they mark the interaction between writer and reader and enable the writer to modulate the force of their arguments and degree of commitment to their claims. Their presence in an encyclopedic text, though, is uncommon and represents a departure from the convention. Such a high frequency of hedges in the Wiki-discourse signifies a modification of the functional and pragmatic aspect of this discourse. Hedges involve the intervention of the writer into the text, since they do not just communicate ideas, but also indicate the writer's attitude to them and to readers (Hyland 2002). As such, they also contribute to the personalization of the discourse and significantly increase the degree of content subjectivity.

C. Interactivity

Discourse personalization can be also observed in the authors' attempts at the establishment of contact with the readers. In this way, the authors not only mark their presence in discourse, but also acknowledge the presence of their readers. The degree of writer-reader interaction in the articles is variable and is achieved by different means. The most frequent are: the use of personal pronouns, especially the *we*- and *you*-constructions as well as direct references to the reader in the forms of the imperative.

a) We-constructions

The use of *we*-constructions, i.e. the use of *we* and first person plural forms of verbs constitutes one of the most frequent means of establishing contact with the readers. *We*-constructions involve the so-called "exclusive *we*", where the pronoun refers either to the author him/herself, or to the authors if the text is an effect of collaboration, and the "inclusive *we*", with the pronoun encompassing not only the author, but also the reader of the message, thus establishing interaction between the two sides. As has been noted in some studies (Chih-Hua 1999, Floettum *et al.* 2007), this combination of pronoun and verb may include not only the reader, but even the whole relevant discourse community, or as Floettum *et al.* (*ibid.*) observe, the whole of mankind. In this way, such types of utterances help to build a community of opinion. Using the first person plural pronoun, the author treats him/herself as one of the readers. The interaction-enhancing role of the "inclusive *we*" is most evident in articles from the field of science. The use of the "inclusive *we*" assigns to such texts a form close to a scientific lecture. The presence of such forms is strongest in the articles explicating certain phenomena and those which provide evidence either validating, or

highlighting the erroneousess of particular theses. Within the analysed material, we encountered 217 such forms in 92 articles (14.42% of the entries). In the extracts below, the authors provide explanations of specific notions. By using the “inclusive *we*”, the authors presuppose the readers’ aptitude to follow a certain line of reasoning (Chih-Hua 1999), e.g.:

Under determined

We can collect the coefficients in a matrix as follows: (...) If we represent each matrix by a single letter, this becomes...

(http://en.wikipedia.org/wiki/Under_determined, 28th June 2007)

Jody Azzouni

There is no standard way to indicate which are which but we can generally tell by context. When someone asks “how many prime numbers are there between 10 and 20?”, we understand it in the non-committing way, and answer “4”. However, when asked “Are there actually any numbers?”, we understand it in the committing way, and he suggests we should answer “No”.

(http://en.wikipedia.org/wiki/Jody_Azzouni, 30th May 2007)

Unit prefixes

The pattern here is that we go backwards from the beginning of the alphabet, starting with z and y, and we follow it up with an alteration of the Greek or Latin for the next number.

(http://en.wikipedia.org/wiki/Non-SI_unit_prefixes, 24th November 2007)

We-forms indicate an interpersonal focus and an involved style in general (Chafe 1982, 1985). By such means, as Floettum *et al.* (2007) note, the authors shorten the distance separating the author from the reader, thus bringing the reader closer. By the use of *we*-constructions, the authors invite the readers to follow the thread of argumentation. The writer guides the reader through an explication but, what is significant, guides him/her toward a preferred interpretation of the notion in question. The authorial stance is explicit in the example below, where the author resorts to an evaluation of the procedures associated with the notions discussed:

UBV system

For example, we can calibrate colour index diagrams UB,BV in the UBV system according to spectral and luminosity classes. Nevertheless, this callibration is not straightforward, because many effects are superimposed in such diagrams: metallicity, binary and multiple stars. The more colours and more narrow passbands in photometric systems we use, the more precisely we can derive star’s class (and, hence, physical parameters). The best are, of course, spectral measurements, but we not always have enough time to get qualitative spectra with high signal-to-noise ratio.

(http://en.wikipedia.org/wiki/UBv_system, 6th April 2007)

In the extracts quoted above, the authors treat themselves as members of a community into which they also draw the reader. In this way, the writer and the reader become equal participants of the communicative act and they are both involved in the solution of a specific problem.

We-constructions may also be used to refer to an area of study, as in the examples presented below, where *we* and *our* refers to the knowledge of the researchers in a particular discipline:

Khwarezmian

Our knowledge of Khwarezmian is limited to its Middle Iranian stage and much like Sogdian, we are not sure of its ancient form.

(http://en.wikipedia.org/wiki/Khwarezmian_language, 4th May 2007)

Uvdal Stave Church

The stave church was originally constructed just after the year 1168, which we know through dendrochronological dating of the pine tree used during the construction. (...) This we know from an archeological excavation that took place during 1978.

(http://en.wikipedia.org/wiki/Uvdal_Stave_Church, 18th August 2007)

In these cases, with the use of *we*-forms, the authors identify themselves with an indefinite group of people, e.g. researchers (Floettum *et al.* 2007). Such constructions help the authors refrain from a specific and straightforward explanation of their role in the text, at the same time, protecting themselves from bearing individual responsibility for the presented content.

b) Direct reference to the reader – you

The explanation of particular processes or concepts proceeds as well through direct reference to the reader, via the use of personal pronouns and personal forms of verbs.

The use of these means of explanation increases the level of informality, since such forms are predominantly characteristic of spoken language or unofficial written forms. As Smith (1985) observes, *you* constitutes the most interactive form among the personal pronouns, due to its explicit acknowledgement of the presence of the reader in the text. Their occurrence in Wikipedia is worth noting, as such forms appear in 24.76% of the articles (158 entries) associated with different domains of knowledge, e.g.:

Upgrade

An “upgrade” is when you replace a product with a newer version of that same product. When you make an upgrade of the same product from one company to the other, you are making a “competitive upgrade”.

(<http://en.wikipedia.org/wiki/Upgrade>, 15th October 2007)

Decompression sickness

When you remove the cap from the bottle, you can clearly hear gas escaping and see bubbles forming in the soda. This is carbon dioxide gas.

(http://en.wikipedia.org/wiki/Decompression_sickness, 3rd June 2007)

While the examples quoted above are more general in reference, the extract presented below gives a different impression. Here, the writer seems to enter into a dialogue with a specific reader, a reader involved in the game, as is the writer himself:

Ufo Aftershock

The game assumes that you took up the Reticulans offer of resettling the most able of humanity in a space station, while allowing the rest to die. (However, if you were a particularly skilled player, you could reject their offer and save the Earth while defending against both the Reticulans and the Biomass). Having lost contact with the Earth, you must find out what happened.

(http://en.wikipedia.org/wiki/UFO:_Aftershock, 3rd March 2007)

The three extracts illustrate the dialogic character of the unconventional Wiki-articles and the authors' propensity towards establishing a closer contact with their readers. Using these pronouns the authors stress their cooperation with readers and introduce a more personal relationship with them.

c) Forms of the imperative

While explaining chosen phenomena, the authors of Wikipedia articles also refer to the readers using imperative forms of verbs. Hyland (2002), drawing on Searle's (1976) taxonomy subsumes such forms under the group of directives, encompassing illocutionary acts "proposing an action intended to be carried out by the reader" (*ibid.*: 216). Directives, depending on their form, can exert different influences on the readers, inducing them to undertake different activities, i.e. physical (involving a real world action), textual (referring readers to a different part of the text or to a different text) and cognitive (directing readers to follow a specific line of argumentation) (*ibid.*).

Physical directives appear in Wiki-articles mainly in the texts which explain the inner workings of notions (96 occurrences in 64 articles). By direct reference to the readers, the authors guide them in carrying out activities or procedures associated with the concept in question:

Boustrophedon transform

If k is odd, then put the number a_k on the right end of the row and fill the row from the right to the left, with every entry being the sum of the number to the right and the number to the upper right. If k is even, then put the number a_k on the left end and fill the row from the left to the right, with every entry being the sum of the number to the left and the number to the upper left. (http://en.wikipedia.org/wiki/Boustrophedon_transform, 4th May 2007)

Underwater diving

Assuming that you can swim on the surface, the main obstacle to diving is likely to be the barrier of immersing your head. To overcome this, try hard to keep your eyes open while under the water. A good exercise is to lie face down in water with a snorkel in your mouth and no mask. Breathe slowly and deeply without using your nose – this will get you used to breathing with mask and snorkel. Don't be afraid of water getting into your eyes. Your eyes and nose will become accustomed to immersion. (...) The crucial step in gaining underwater mobility is adopting a suitable posture. To do this, first try to reach an object on the floor of the pool that is within our depth. It will be difficult to reach from an upright posture. To get your hands to the object, jump up, bend your body well forward, throw your feet in the air and try to reach the object, head foremost.

(http://en.wikipedia.org/wiki/Underwater_diving, 6th January 2007)

It is interesting that despite bibliographical lists and references inserted into the articles, the authors resort as well to textual directives, advising readers directly and referring them to other sources for consultation, as in the following examples:

Ummah

See Dhimmi for a full discussion of this concept.

(<http://en.wikipedia.org/wiki/Ummah>, 14th January 2007)

Umuahia

For more information about Umuahia and it's people, visit Umuahia Online – www.umuahia.org (<http://en.wikipedia.org/wiki/Umuahia>, 10th April 2007)

Ube Yamaguchi

See the following page [1] for more information and images of some current exhibits.
(http://en.wikipedia.org/wiki/Ube,_Yamaguchi, 21st August 2007)

Textual types of directives occurred in 18.33% of the articles (117 entries). Still, it is mainly the cognitive type of directives that is employed by Wikipedia authors, present in 32.60% of the entries (306 occurrences in 208 articles). These are realized mainly by the verbs *consider* and *note*. For example:

Lake Toba

To give an idea of its magnitude, consider that although the eruption took place in Indonesia, it deposited an ash layer approximately 15cm thick over the entire Indian subcontinent.
(http://en.wikipedia.org/wiki/Lake_toba, 9th June 2007)

Uniform continuous function

Every uniformly continuous function is continuous, but the converse is not true. Consider for instance the function $f(x)=1/x$ with the positive real numbers. This function is continuous, but not uniformly continuous, since as x approaches 0, the changes in $f(x)$ grow beyond any bound.
(http://en.wikipedia.org/wiki/Uniform_continuous_function, 27th June 2007)

The *note*-forms of the imperative appear especially in digressions, which turn the readers' attention to different aspects of a particular problem, e.g.:

Uralic peoples

Note that in any event the classification "Uralic" addresses matters of language, which do not necessarily correlate with divisions of ethnicity or even of specific culture.
(http://en.wikipedia.org/wiki/Uralic_peoples, 23rd September 2007)

Uniformitarianism

Note, however, that many "catastrophic" events are perfectly compatible with uniformitarianism.
([http://en.wikipedia.org/wiki/Uniformitarianism_\(science\)](http://en.wikipedia.org/wiki/Uniformitarianism_(science)), 4th October 2007)

Upper-triangular matrix

However, note that the product of a lower triangular with an upper triangular matrix does not preserve triangularity.
(http://en.wikipedia.org/wiki/Upper-triangular_matrix, 1st October 2007)

In these instances, the imperatives introduce an effect of interaction, and emphasize important points, as they instruct the reader on how a process or a notion is to be perceived.

In 13.16% of the entries (84 articles), the imperatives appeared in metatextual comments, which constitute a further example of the interaction between author and reader. Such comments concern mainly the form of content presentation, where authors provide readers with advice on how their articles should be approached, e.g.:

Utaibito Hane

Please note, names are given here in Western personal name – surname order.
(http://en.wikipedia.org/wiki/Utaibito_Hane, 19th April 2007)

Barbireau

Note that this is not a list of all artists who have ever been members of the Roman Catholic Church. Please do not add entries here without providing support for those artists having specifically Roman Catholic artistic themes in their art.
(http://en.wikipedia.org/wiki/List_of_Roman_Catholic_Church_artists, 15th March 2007)

In contrast to the previous forms establishing interaction, directives emphasize the expertise of the author of the text (Hyland 2002: 217).

As can be observed in the examples presented above, the use of pronouns and forms of the imperative reveals the authors' relationship with the readers. The texts are aimed at non-specialists with the goal of transferring information in a simplified manner, adjusted to the capabilities of the potential receivers. As the examples show, such a way of text formulation makes it more evocative and suggestive to the readers. Interactivity-enhancing elements recognize readers as participants within the discourse, help stimulate them and focus their attention on different aspects of the issues discussed.

6.1.2.3.2. Informality

The informality of the discourse reflected in the textual and syntactic patterns has been already described in the previous sections, but the presence of colloquial expressions is manifested also in the lexical plan of the free encyclopedia, which has been partly illustrated in the previous subsection on the evaluative and subjective vocabulary. However, it can be observed also in further areas of the lexical stratum. Within the colloquial lexis, the most frequent are content words as well as idiomatic expressions and contractions.

Within the selected articles, the colloquial nouns number 2,934 with 1,829 adjectives:

150

Up in the Attic

Its considered a commercial flop in the US.

(http://en.wikipedia.org/wiki/Up_in_the_Atlic, 16th July 2007)

Ursula Paerson Bellah

Born Usula Gadischke in 1929, the girl who went by "Ulli" had a tough time growing up.

(http://en.wikipedia.org/wiki/Ursula_Pearson, 1st October 2007)

Udinese Calcio

It has a good number of fans in Friuli and surrounding areas, and it is sometimes seen as the best symbol of Friulian pride.

(http://en.wikipedia.org/wiki/Udinese_Calcio, 28th September 2007)

Up against it

Many critics stated that this was a bad move.

(http://en.wikipedia.org/wiki/Up_Against_It, 2nd July 2007)

Adverbs (1229):

Uranus

He was simply the sky.

([http://en.wikipedia.org/wiki/Uranus_\(mythology\)](http://en.wikipedia.org/wiki/Uranus_(mythology)), 12th November 2007)

Urith

Urith is a fairly obscure figure.

(<http://en.wikipedia.org/wiki/Urith>, 15th January 2007)

Uvdal Stave Church

Two scary half masks are quite visible on the poles of the chancel.

(http://en.wikipedia.org/wiki/Uvdal_Stave_Church, 1st July 2007)

Colloquial verbs (2534):

Ufa

The city began to be called Ufa, meaning “small” in Turkic, by locals and the name stuck.
(<http://en.wikipedia.org/wiki/Ufa>, 23rd October 2007)

Ubastet

In Sandman, Basat appears as a friend and confidant of Dream, at one point openly flirting with him although nothing comes of it.
(<http://en.wikipedia.org/wiki/Ueuecoyotl>, 6th April 2007)

Under the Pink

The original tracklisting included the b-side “Honey,” which was kicked off at the last moment in favour of “The Wrong Band”.
(http://en.wikipedia.org/wiki/Under_The_Pink, 10th March 2007)

Phraseologisms and idiomatic expressions (896):

UEFA

UEFA as a representative of the national associations, has had a number of bruising clashes with the European Commission.
(<http://en.wikipedia.org/wiki/Uefa>, 6th October 2007)

CityBird

Thomas Cook toyed with the idea of buying them out of bankruptcy, but later pulled out.
(http://en.wikipedia.org/wiki/City_Bird, 8th March 2007)

Udo Bolts

He distinguished himself as a courageous rider who liked to get into breakaways.
(http://en.wikipedia.org/wiki/Udo_Bolts, 24th May 2007)

Uniklubi

He is the part and parcel of Uniklubi ever since.
(<http://en.wikipedia.org/wiki/Uniklubi>, 23rd June 2007)

Umberto Bossi

When the scandals of Tangentopoli were unveiled from 1992 on, Bossi rode the wave, presenting himself as the new man in politics, set out to sweep away corruption and incompetence.
(http://en.wikipedia.org/wiki/Umberto_Bossi, 19th February 2007)

Uwe Reinders

After handing up his boots he started his managerial career at then 2nd Bundesliga side Eintracht Braunschweig in 1988-1989.
(http://en.wikipedia.org/wiki/Uwe_Reinders, 1st March 2007)

The informality of Wiki-articles is also reflected in a high number of contractions (1309):

Uriah Grant

Grant hasn't fought since the loss to Castillo.
(http://en.wikipedia.org/wiki/Uriah_Grant, 3rd May 2007)

Uriankhai

There isn't a clear ethnic delineation for the application of the name “Uriankhai”.
(<http://en.wikipedia.org/wiki/Uriankhai>, 23rd September 2007)

Ulpia Marciana

Marciana first didn't accept this, but her sister in law insisted on her to have this title.
(http://en.wikipedia.org/wiki/Ulpia_Marciana, 25th March 2007)

The orality of these Wiki-articles may be as well observed in the presence of deictic expressions in the articles. Especially frequent is the use of place and time adverbials, which report events as being in progress at the time of writing (overall 348). Since these means are used to refer directly to the immediate physical and temporal context, they remain relevant only for a specific period of time, which undermines the universal validity of an article. For instance:

Ulteo

Ulteo's website previously estimated that a first beta would be released in October or November 2006. The homepage now talks about 'soon'.
(<http://en.wikipedia.org/wiki/Ulteo>, 23rd January 2007)

Up or down vote

To this end, many Republican-affiliated websites and political action committees have recently urged voters to demand of their Senators and Representatives an "up or down vote" on various issues.
(http://en.wikipedia.org/wiki/Up_or_down_vote, 18th June 2007)

Ulyanovskaya mine disaster

It has also been announced that 60 coal mines in the surrounding area were inspected for similar violations shortly after the disaster, and that the entirety of the nation's mines are to be inspected over the coming weeks.
(http://en.wikipedia.org/wiki/Ulyanovskaya_mine_disaster, 31st May 2007)

Wiki-articles resemble a spoken conversation with the informal means of expression, such as those quoted above, performing an expressive function, as they present feelings and attitudes and constitute a means of a subjective evaluation of the content (Moon 1987). The colloquial means of expression fulfil a phatic function, facilitating contact with the readers, attracting and keeping their attention plus enhancing the attractiveness of the message. Yet, one should differentiate between instances in which colloquial means are used to explain an intricate notion, and instances which rather signify the author's lexical poverty.

6.1.2.3.3. Terminology

Wiki-articles exhibit different approaches with regard to the use of specialized terminology in the definitions, which reflects a lack of uniformization in this respect.

A common (154 instances) means of introducing specialized terminology is its juxtaposition with an explanation or a synonymous phrase, given in the form of a periphrasis or an explicative metatextual commentary enclosed in brackets, which serve as explanations of scientific or unclear terms used in the text. Such comments constitute an attempt at avoiding the hermetism of a scientific style and prove that the authors use specific scientific terminology consciously and purposefully:

U-Boat

When using magnetic torpedoes, the commander would know the target ship's draft (how low its keel sits in the water) and set the torpedo's depth.
(<http://en.wikipedia.org/wiki/U-boat>, 6th April 2007)

UBV System

A few dwarf (that is, main sequence) carbon stars are known.
(http://en.wikipedia.org/wiki/Ubv_system, 6th April 2007)

By such metatextual comments, the authors illustrate their linguistic competence and awareness of the difference between scientific and colloquial register, e.g.:

Umbilicus

In humans, the scar can appear as a depression (sometimes colloquially referred to as an *innie*) or as a protrusion (*outie*).
(<http://en.wikipedia.org/wiki/Navel>, 18th October 2007)

Yet, as far as this approach is concerned, Wiki-articles exhibit a lack of standardization and homogeneity. Within one group of the articles it is the explication of the specialized term that is enclosed in brackets, e.g.:

Boustrophedon transform

The transformed sequence is computed by filling a traingle in boustrophedon (zigzag) manner.
(http://en.wikipedia.org/wiki/Boustrophedon_transform, 4th May 2007)

Budget

In a personal or family budget all sources of income (inflows) are identified and expenses (outflows) are planned with the intent of matching outflows and inflows (making ends meet).
(<http://en.wikipedia.org/wiki/Budget>, 21st September 2007)

In others, the parentheses enclose the terms which follow the explanation, or both techniques may be found in the same articles, e.g.:

Uwaisyaan

In the science of spirituality of Islam (Tasawwuf) the Uwaisi Transmission occurs when the spirits of righteous believers (saliheen, awliya) meet in the world called 'alam al-arwaah (the world of spirits) which is beyond 'alam al-ajsam (the material plane).
(<http://en.wikipedia.org/wiki/Uwaisi>, 28th January 2007)

Metalinguistic cues may have the form of an authorial comment, a periphrasis to guarantee a more exact explanation of the phenomenon. Such comments show that the author perceives the suggested periphrasis as an inadequate explanation of the original term. This is signalled by such expressions as *it is thought of as* or *roughly*, which were encountered in 13.63% of the articles (96 instances in 87 entries), e.g.:

Uronic acid

An uronic is best thought of as a sugar in which the terminal carbon's hydroxyl function has been oxidized to a carboxylic acid.
(http://en.wikipedia.org/wiki/Uronic_acid, 4th June 2007)

Ueshiba Kisshomaru

Kisshomaru took on the mantle of 'doshu', which roughly translates as 'keeper of the way'.
(http://en.wikipedia.org/wiki/Ueshiba_Kisshomaru, 8th July 2007)

Within 7.6% of the articles (49 entries), several of the above-mentioned devices were brought together in one definition, as in the following example, where the author relates the popular term for a species to its scientific name using the conjunction *or* and deploying inverted commas for the commonly used term:

Urodidae

Urodidae or “false burnet moths” is a family of insects.
(<http://en.wikipedia.org/wiki/Urodidae>, 10th August 2007)

Uperoleia

Uperoleia is a genus of frogs, native to Australia in the family Myobatrachidae. These are small squat frog, more commonly known as “toadlets”. They have granular skin, often with a pair of raised glands behind each eye, or on the flanks. They have bumpy, rough skin giving them the appearance of a small toad hence the name “toadlet”.
(<http://en.wikipedia.org/wiki/Uperoleia>, 21st September 2007)

As the last of the extracts presented above shows, the informative purpose of the definition may be also strengthened by explicative remarks provided by the authors, referring to etymological properties of the term (*hence the name*).

All of the aforementioned techniques of introducing specialized terminology have the purpose of facilitating the comprehension of definitions. To achieve this aim, the authors frequently (173 instances) resort to the use of figurative language and similes associated mainly with everyday experiences, e.g.:

Urinary system

The urinary bladder is a hollow muscular organ shaped like a balloon.
(http://en.wikipedia.org/wiki/Urinary_system, 30th October 2007)

Ubiquitin-proteasome pathway

The proteasome is a large barrel-like complex.
(http://en.wikipedia.org/wiki/Ubiquitin-proteasome_pathway, 6th April 2007)

Though, as marked above, the majority of the texts resort to the use of concrete vocabulary, a mention must be made also of entries representing a different approach, namely articles overloaded with discipline-specific terminology, whose content comprises lexemes which have abstract and/or foreign meanings. Within the analysed material, such articles amounted to 9.24% (59 entries), e.g.

Urocortin

Urocortin is a potent anorexigenic peptide of 40 amino acids that induces fed-like motor activity when administered centrally or peripherally in fasted animals. Urocortin belongs to the corticotropin-releasing factor (CRF) family that includes CRF, urotensin I, sauvagine, urocortin II and urocortin III. Urocortin is also a potent and long-lasting hypotensive agent and increases coronary blood flow.
(<http://en.wikipedia.org/wiki/Urocortin>, 8th June 2007)

Uranium

Uranium-thorium dating, also called thorium-230 dating, uranium-series disequilibrium dating or uranium-series dating, is a radiometric dating technique commonly used to determine the age of carbonate materials such as speleothem or coral. Unlike other commonly used radiometric dating techniques such as rubidium-strontium or uranium-lead dating, the uranium-thorium technique does not measure accumulation of a stable end-member decay product, instead cal-

culating an age from the degree to which equilibrium has been restored between the radioactive isotope thorium-230 and its radioactive parent uranium-234 within a sample.

(http://en.wikipedia.org/wiki/Uranium-thorium_dating, 11th April 2007)

Urodus parvula

The “Bumelia Webworm Moth” *Urodus parvula* is recorded on Lauraceae: (avocado=*Persea*), Fagaceae (*Quercus*), Sapotaceae (*Sideroxylon*) and Erythroxylaceae: *Erythroxylum*. *Urodus Parvula* has also been reared on Rutaceae (*Citrus*) and Malvaceae (*Hibiscus*). *W. asperipunctella* has in North America been reared from aspen *Populus tremuloides* and also *Salix* in Europe.

The pupa is contained in an open-mesh cocoon, which can be bright orange in colour.

(<http://en.wikipedia.org/wiki/Urodidae>, 10th August 2007)

An exaggerated use of specialized terms may lead to an informative overload in the message, which may hinder its intelligibility, as in the following article, where the accumulation of discipline specific terms makes the entry quite incomprehensible to a layman. Even if the authors resort to techniques of explication, e.g. provide an explanation in brackets, the definition, dense with specialized terms, does not necessarily result in greater clarity of meaning:

Upington disease

The disease is characterized by Perthes-like pelvic anomalies (premature closure of the capital femoral epiphyses and widened femoral necks with flattened femoral heads), enchondromata and ecchondromata.

(http://en.wikipedia.org/wiki/Upington_disease, 30th September 2007)

Variation in the use of specialized terminology reflects the different aims as well as the different potential audience the authors have in mind while constructing their articles. A lack of definitions in the selected scientific articles presupposes that an audience comprises a disciplinary community for whom the denotation of the specialized terms is clear. In such cases, the knowledge of specific terms is taken for granted by the authors and no further clarification is given. On the contrary, in the cases where the authors write for a lay audience, the percentage of definitions as well as the variety of clarification techniques is higher. It is mainly the latter approach that is applicable in Wikipedia, as is confirmed by the treatment of terminology. By means of the above mentioned techniques, the authors try to decrease the scientific character of the description.

Such a wide differentiation of approaches towards the use of specialized terminology and types of defining techniques contribute to a general impression of heterogeneity in Wikipedia, illustrated in the previous section describing its textual characteristic, and again manifests a lack of standardization of the techniques of description.

6.1.2.3.4. Inverted commas

As observed in the previous section, inverted commas are used to introduce common equivalents of specialized terms into definitions. In such cases, inverted commas mark the words or phrases which stand out from the text as instances of a different register, usually colloquial, which contrast with the official character of the text. The writers also tend to use colloquial words in inverted commas in order to provide a readable and simple explanation of scientific processes (78 instances in 638 articles).

Colloquial words in such contexts are often used metaphorically, and refer to concrete, common phenomena to give readers an illustrative definition of specific concepts:

Ubiquitin-proteasome pathway

The outer two rings each contain seven subunits whose function is to maintain a “gate” through which proteins enter the barrel.

(http://en.wikipedia.org/wiki/Ubiquitin-proteasome_pathway, 6th April 2007)

Uperoleia

There are two distinct types of calls Uperoleia species make, either a “click” or a “squelch”.

(<http://en.wikipedia.org/wiki/Uperoleia>, 21st September 2007)

The authors very often (83 occurrences) put colloquial phrases into inverted commas. Most frequently, such terms comprise commonly used set phrases or idiomatic expressions:

Triad

Both Ixion and Skydive were “in” the computer business. (...) They offered HK a “deal” – if HK would keep supplying Traid with originals, Traid would not spread the cracked games in Sweden.

(http://en.wikipedia.org/wiki/Triad_C64, 3rd June 2007)

Under the Bridge

It left out the final verse, which includes the only drug reference in the song – likely done to maintain the girl group’s “clean” image.

(http://en.wikipedia.org/wiki/Under_the_Bridge, 24th September 2007)

Uber Geek

Behavioral traits that one sports as an indication of being “in the know” and “out of the mainstream”.

(http://en.wikipedia.org/wiki/Uber_geek, 23rd September 2007)

In 7.6% of the articles (49 out of 638), inverted commas were introduced with the purpose of indicating that the words do not represent their literal, obvious meaning, but rather an approximation, or an ironic and humorous sense, e.g.:

Upsala Simsallskap

The same year the society was founded, Svanberg arranged a mock “graduation ceremony”, where he and some of his colleagues awarded themselves the master’s degree and their pupils the degree of bachelor in swimming. The simpromotion and these “degrees” were taken over by other Swedish swimming clubs and are still used in Swedish swimming schools.

(http://en.wikipedia.org/wiki/Upsala_Simsallskap, 15th March 2007)

Brocius

Brocius is described by contemporary Billy Breakenridge in his book Helldorado as being the most deadly pistol shot of the Cow-boys, able to hit running jackrabbits, shoot out candle flames without breaking the candles or lantern holders, and able to shoot quarters from between the fingers of “volunteers”. When drunk he used gunfire to make a preacher “dance” during a sermon.

(http://en.wikipedia.org/wiki/William_Brocius, 17th April 2007)

The non-literal sense of the words used is clearly evident in the instances given below:

Umar al-Bashir

He was later “elected” president in a showcase national election in 1996.

In 1999, al-Bashir and the Parliament made a law which allowed limited political “associations” in opposition to al-Bashir and his supporters to be formed, although these groups failed to gain any significant access to governmental power.
(http://en.wikipedia.org/wiki/Umar_al-Bashir, 6th January 2007)

Corralito

Some international banks negotiated with the Argentine governments to receive compensation bonds for the “missing” money, which in a large proportion had never really left their banks.
(<http://en.wikipedia.org/wiki/Corralito>, 18th September 2007)

U-boat

Before the start of World War II, Germany started rebuilding U-boats and training crews, hiding these activities as “research”.
(<http://en.wikipedia.org/wiki/U-boat>, 6th April 2007)

In these instances, the use of inverted commas is not aimed at exposing informality, but rather is to suggest that the words should not be accepted verbatim. In such a way, the authors distance themselves from the words they are using (Carter and McCarthy 2006) and, at the same time, show that they do not agree with the meanings intended by these words. In these cases, inverted commas also seem to function as a means of attracting the users’ attention.

It should also be mentioned that there are unnecessary uses of inverted commas (46 instances), as in the example below, where the application of quotation marks seems redundant:

UEFA Champions League

As a result, the system was restructured to force “weaker” national champions to qualify for the group stages, while other “stronger” national runners-up would automatically get places.
(http://en.wikipedia.org/wiki/UEFA_Champions_League, 16th December 2007)

Ultra Violet – Dufresne

After meeting with Salvador Dali, she became his “muse” and pupil and then became an artist in her won right.
([http://en.wikipedia.org/wiki/Ultra_Violet_\(Isabelle_Collin_Dufresne\)](http://en.wikipedia.org/wiki/Ultra_Violet_(Isabelle_Collin_Dufresne)), 28th August 2007)

In these extracts the use of inverted commas seems to be superfluous, as the words do not indicate an uncommon usage of the words in question and neither do they stand out from the text as instances of different register.

The frequency of subjective, interactive and colloquial forms and the emphasis placed on the techniques directed at the achievement of a persuasive influence upon the reader proves that Wikipedia articles are not simple statements of information comparable to traditional printed encyclopedic entries. Through the introduction of the above-mentioned elements, the Wiki-discourse departs from the conventional encyclopedic tone, leaning towards popular-scientific, educational or journalistic texts and even towards the tone of a colloquial conversation.

In conclusion, unconventional Wiki-articles manifest different degrees of departure from the standard properties of encyclopedic discourse on all levels of discourse. The primary difference between conventional and unconventional Wiki-articles concerns a transition on the continuum leading from formality and detachment towards informality and involvement, and from homogenization towards hybridization and syncretism.

6.1.3. Anti-encyclopedic articles

The analysed collection of 986 Wiki-articles allows us to single out yet a further small group of texts, leaning towards the opposite pole of the continuum, which is worth mentioning. The articles comprising nearly 12% (114 articles – 11.77%) of the whole material constitute examples of anti-encyclopedic forms, since they exhibit features which disentitle them from being instances of encyclopedic articles due to their considerable departure from the convention. These articles illustrate various types of deviations which impinge on the basic, communicative function of the articles. The main criterion which is violated by the anti-encyclopedic articles is the requirement concerning intelligibility, precision, coherence and linguistic correctness of the message.

6.1.3.1. Unintelligibility

The violation of the intelligibility criterion constitutes a consequence of an insufficient mastery of the language. This group of articles (39 articles, i.e. 4.02%) comprises mainly the entries written presumably by non-native speakers of English who do not possess the appropriate language skills, which would enable them to construct a text of this genre. For instance:

Urakami clan

Urakami clan a Japanese clan that primarily consisted at Bizen Province throughout the course of the Sengoku Period of the 16th century. One major turn of events that took place for the Urakami was through the death of Urakami Norimune folloowing the year of 1502. This led to Matsuda Motokatsu to take the chance to lead a full scale assault on the domain of the Urakami. One major retainer of the Urakami was UkitaYoshiie, in which he led 300 troops in the attack against the Matsuda clan (in which at the time the Urakami were allied with the Ukita). The Urakami managed a somewhat balanced course throughout the Sengoku Period following this event.

(http://en.wikipedia.org/wiki/Urakami_clan, 18th April 2007)

Uzbek cuisine

Uzbek cuisine is influenced by the general agriculture as in most nations. There is a good deal of grain in the nation so breads and noodles are of importance. In discussing the cuisine of Afghanistan, The Village Voice mentions the “noodle-rich cuisine of Uzbekistan.” In meats mutton is popular and is a part of various dishes in the country. The most popular beverage might be green tea, and teahouses are of cultural importance. Alcohol might be less popular in the West, but wine is comparatively popular for a Muslim nation as Uzbekistan is largely secular. Samarkand is perhaps the city best known for a having winery.

(http://en.wikipedia.org/wiki/Uzbek_cuisine, 21st March 2007)

Ulanów

Torches lit as the wooden hand made rafts continue and are not allowed to return to shore unless the torches are burnt out. There is also a boat containing the mayor of the town with a large cannon that is shot when the ceremony is over which is then followed by fireworks.

(<http://en.wikipedia.org/wiki/Ulan%C3%B3w>, 17th August 2007)

A closer look into the content of the articles quoted above proves that the non-encyclopedic character of these messages constitutes a consequence of their authors’

language incompetence. The authors of these entries may have an awareness of the genre and of the purpose of the article, yet their language incompetence prevents them from transmitting the content in a correct and intelligible manner. What is especially striking in these texts is a specific ineptitude to transfer the ideas into a readable textual format. The awareness of the informative purpose of the genre is reflected in the attempt the authors make to pass on chosen facts concerning the Urakami clan, the culinary traditions of Uzbekistan and a festival in Ulanów. Yet, the improper formulation of the message prevents the achievement of this purpose. As far as the textual level of discourse is concerned, the messages are somewhat incoherent, the entries contain incorrect grammatical constructions and inappropriate vocabulary items. Thus, the informative aim of the articles can be only attained provided the readers put some effort into deciphering the message and deriving the necessary information from the content themselves.

A further instance of unintelligibility, and thus infringement of genre properties, may be exemplified in the entries below. These extracts illustrate an instance of contamination of the content with foreign, incomprehensible elements. In these cases, heterogeneity leads to the effect of complete meaninglessness and incomprehensibility of the articles:

Mannarkkad

Mannarkkad is a small town in Palakkad district of Kerala state, south India. Silent Valley is only 66 km from Mannarkkad.

Mannarkkad still keeps the terrific memories of Tippu Sulthan and Mophla riot. In 1921, hundreds of Hindus lost life and thousands of Hindus converted to Islam forcefully. The town has lost its cultural diversity as sandal smugglers and terror outfits emerge. the former ruler/ desavazhi of mannarkkad is known as “mannurkkadu muppil nair”. he was a “samantha” under “valluvakonathiri” of valluvanadu. the famous novel of malayattoor ramakrishnan “ponni” is based on the adivasis of mannarkkad. sandalwood smuggler veerappan used ‘shiruvanni forests’ of mannarkkadu muppil nairs as his shelter.

(<http://en.wikipedia.org/wiki/Mannarkkad>, 28th May 2007)

Uthangarai

Uthangarai is a panchayat town in krishnagiri district in the Indian state of Tamil Nadu.

MRS.RAJAMMAL RAJU WARD MEMBER OF TOWN PANCHAYATH

R.ARTHANARI WAS LIVED IN UTHANGARAI.

HE HAD FOUR SONS A.BALAJI, A.RAJU , A.MUTHU , A.ARUMUGAM

(<http://en.wikipedia.org/wiki/Uthangarai>, 8th September 2007)

The articles exhibit traces of encyclopedic content, which can be seen in the introductory sentences containing general explications of the headwords. Yet, what follows is a radical change in the stylistic layer of the entries and the manner of content description. While in the first instance the additional content is, to an extent, associated with the rest of the article, though incomprehensible to a reader not knowing Hindi, in the second example, the message written in capital letters constitutes an explicit example of an introduction of senseless and non-encyclopedic content. These extracts exemplify not only the authors' low language proficiency, but also a lack of discourse competence and genre awareness. An infringement on the intertextual sur-

face of discourse in these cases contributes to a serious impairment of the functional value of the text, so that it becomes virtually impossible to assess the communicative aim of the message.

6.1.3.2. Imprecision

One of the most important requirements with which a good definition should comply is precision and a sufficiency of the information provided. Deviant definitions in this respect concern entries which are imprecise and as such do not attain the level of accuracy needed to understand the concept defined. A departure from the norm is exemplified by incomplete, circular and ill-ordered definitions.

A lack of precision is clearly evident in incomplete definitions, that is such entries in which the headword is not accompanied by a sufficient and meaningful explication. This problem concerns in particular the so-called “stubs”, i.e. short introductory entries, which wait to be expanded. Under each stub, there is a notation inviting an editor to extend the content of the article. It should be noted that some articles which have the status of a stub are still meaningful and provide readers with basic information on the subject. Yet, within the analysed material, 1.13% of the articles (11 entries) do not attain even the minimum level of informativity, as they do not in any way explicate the headword of the entry. For instance:

Upper Knowle

Upper Knowle is located in south Bristol.

(http://en.wikipedia.org/wiki/Upper_Knowle, 14th May 2007)

Which does not tell the reader what *Upper Knowle* actually is.

A further feature impinging on preciseness entails the frequent circularity of definitions observed in 2.16% of the entries (21 articles). Circularity consists in the repetition of the defined term in the same or only a slightly modified form. This leads to a roundabout definition, where, in effect, the meaning of the term is not properly explained. For instance:

Compact manifold

A compact manifold is a manifold that is compact as a space.

(http://en.wikipedia.org/wiki/Compact_manifold, 16th October 2007)

In the example above, the author does not provide an explication of the meaning of the term *manifold*. Such a definition does not have much informative value, as it leaves the reader without any meaningful clarification of the issue in question. Similar examples of roundabout definitions are evident in the following articles:

Unconditional surrender

Unconditional surrender is a surrender without conditions.

(http://en.wikipedia.org/wiki/Unconditional_surrender, 8th February 2007)

Udmurts

The Udmurts are a people who speak the Udmurt Language.

(<http://en.wikipedia.org/wiki/Udmurts>, 21st May 2007)

Yet, not only circularity of information undermines clarity, but also erroneous structuring of the explanation. As described above, intelligibility of information can be achieved by a proper form of the message, where the most important facts are presented in the initial parts of the definition, following directly the concept which is being clarified. Certain Wikipedia articles (17 entries, i.e. 1.75%) exhibit a departure from this convention, and thus impair the precision of the message. Such practice is exemplified in the following extract, where the information fundamental to the comprehension of the subject is moved away from the headword:

Umar Mokhtar

Umar Mokhtar was from the tribe of Mnifa, born in a small village called Janzour located in the eastern part of Barqa – not to be confused with the city of western Libya called Janzour which is more well known. He was the leader of the resistance movement against the Italian military occupation of Libya for more than twenty years.

(http://en.wikipedia.org/wiki/Omar_Mukhtar, 20th June 2007)

The impression of the incompleteness of the information is also produced by other means, which include adverbial phrases such as *and so on* (57 occurrences in the analysed group) or *and more* (23 instances), amongst others, for instance:

Udmurt language

Udmurt is an agglutinating language, it uses affixes to explain possession and so on. Many loanwords are from the Tatar language, as well as phonetics and so on.

(http://en.wikipedia.org/wiki/Udmurt_language, 9th September 2007)

Udmurt

Udmurtia's natural resources include oil, peat, and more.

(http://en.wikipedia.org/wiki/Udmurt_Republic, 13th September 2007)

Umiastowski family

The Umiastowski family line is a Polish-German line, as at one point in time the Umiastowski settled in an area now belonging to Latvia and Estonia and tended to marry women of German origin.

(http://en.wikipedia.org/wiki/Umiastowski_family, 15th October 2007)

As the examples illustrate, the authors of Wikipedia articles do not always aim at attaining a high level of exactness in their explications.

6.1.3.3. Incoherence

Textual coherence constitutes a significant condition determining the clarity of a message. The degree of coherence may be increased by the correct deployment of cohesive devices, which point to the relationship between particular propositions.

While a substantial proportion of Wiki-articles exhibit a coherent textual framework, it is worth noting that there are frequent instances of a certain ineptitude in the employment of referential devices resulting in the infringement of textual coherence, which exerts an unfavourable effect on the overall degree of readability and informativity of the articles, thus impairing the communicative aim of the texts.

a) Unclear reference

Problems concerning reference devices in Wiki-discourse entail their erroneous use, which not only undermines the coherence of the articles but, more importantly, might lead to the readers' misinterpretation of the message. We identified 19 instances of such mistakes.

In the first example, the lack of clarity refers to the pronoun *both*. It is not clear whether it encompasses Uqba and Muhammad, or Uqba and Ubay:

Uqba ibn Abu Mu'ayt

On one occasion, Uqba spit on Muhammad's face at the incitement of his friend Ubay ibn Khalaf and then both died in 624.

(http://en.wikipedia.org/wiki/Uqba_ibn_Abu_Mu%27ayt, 19th August 2007)

In the following case, the ambiguity concerns the patient of the stress caused by forcing Umbreon to walk:

Umbreon

Umbreon are also said to evolve as a result of love for their trainers and the moonlight; forcing one to walk during the day can also cause stress.

(<http://en.wikipedia.org/wiki/Umbreon#Umbreon>, 21st November 2007)

b) Lack of concord

A further feature impairing the clarity of the articles on the textual level consists of the lack of concord between sentences. Most commonly, the inconsistency concerns the tense of description (23 instances) and reader-reference mode (18 instances), e.g.:

Oxyd

To get to the next landscape, the player must open all the Oxyds on the current landscape, which is done by touching them, but they will only stay open if you touch Oxyds of the same color in sequence.

(<http://en.wikipedia.org/wiki/Oxyd>, 2nd January 2007)

In this article, the author moves from the use of the impersonal form of *the player* or *one*, to a direct reference to the reader in the form of *you*. This change entails also a modification in the degree of formality, with the pronoun *you* introducing a sign of informality.

c) Conjunction errors

Wiki-authors seem to encounter problems also with applying conjunctions correctly. Within the analysed material of 96 entries, we encountered 35 instances of conjunction errors.

In the first instance, the clause introduced by *therefore* does not seem to logically result from the information presented in the preceding clause:

Utaibito Hane

Is a Japanese male folk band. Its members are Shinichi Yasuoka and Tetsuro Honda. In 1998 they met each other and formed Hane. Afterwards they changed its name Uaibito Hane. Performing many times on the street in Fukuoka, they succeeded in winning popularity. Therefore they came out with the song "Chiisana Hoshi no Chiisana Tabibito" in 1999. Their song "Mi Title" is the first ending of the anime Hungry Heart: Wild Striker.

(http://en.wikipedia.org/wiki/Utaibito_Hane, 19th April 2007)

By analogy, in the extract below, the conjunction *and* does not seem to rightly fit the context. The clause as it stands provides an unclear presupposition stating that the decline of the party constituted a consequence of its union with the Italian Liberal Party. Thus, it is not clear, whether this union constituted an immediate reason for the decrease in popularity, or was conditioned by different factors:

Uomo Qualunque Front

After a remarkable result in the 1946 general election (4.4%), the party joined forces with the Italian Liberal Party in the National Block and its decline was unstoppable.
(http://en.wikipedia.org/wiki/Uomo_Qualunque_Front, 6th April 2007)

Similarly, in the extract below the author thematizes the fact that Eco's father had twelve siblings, which thus seems to be the direct reason why he urged his son to pursue legal studies:

Umberto Eco

His father was the son of a family with thirteen children, and urged him to become a lawyer, but he entered the University of Turin in order to take up medieval philosophy and literature.
(http://en.wikipedia.org/wiki/Umberto_Eco, 24th August 2007)

The repetition of *and* to connect independent clauses gives the effect of "run-on" sentences, which could be rephrased into more fully embedded syntactic structures:

Umkirch

Sigillat shards and coins were found and the church was built onto a Roman building, as a part of its wall was found.
(<http://en.wikipedia.org/wiki/Umkirch>, 19th March 2007)

In the example below, on the other hand, *and* seems to be used erroneously in the place where *but* is more fitting due to the contrasting meanings of the two clauses:

Babrius

There is no mention of Babrius in ancient writers before the beginning of the 3rd century AD, and his language and style seem to show that he belonged to that period.
(<http://en.wikipedia.org/wiki/Babrius>, 11th June 2007)

Due to the mistakes enumerated above, the defects and specific ineptness in the composition of the articles cause them to become unclear, illogical and imprecise. In consequence, the texts cease to perform their primary function.

6.1.3.4. Language mistakes

Deviations from the norm encompass additionally incorrect language and a lack of stringent editing of the articles. Wikipedia articles are frequently erroneous and not communicative due to numerous syntactic, lexical and punctuation errors.

a) Syntactic mistakes

Syntactic structures may be impaired by flaws in word order encompassing ellipsis or redundancy of reference (46):

Uppland inscription

Gylla was Tolir's wife. Hakon is believed to the reigning king Haakon the Red.
(http://en.wikipedia.org/wiki/Uppland_Runic_Inscription_11, 30th June 2007)

Walkover

It can however be interpreted by critics of the fraction the walkover is awarded to as a suspicious sign of electoral fraud or gerrymandering to prevent other candidates from participating. (<http://en.wikipedia.org/wiki/Walkover>, 26th April 2007)

Past:Present:Future

It is expected it to fall in chart positions or even disappear out of the chart due the fact that Christmas is over and less people will be buying the single or it to have less airplay than before. (<http://en.wikipedia.org/wiki/Past:Present:Future>, 24th February 2007)

Syntactic flaws entail as well the construction of long compound clauses, which may impair the intelligibility of the entry:

Phelesmon

Phelesmon is rumored to be the Ultimate form of Impmon, however, this fact can not be concluded at this time, due to the fact that Impmon only Warp Digivolved during Digimon Tamers to Beelzemon, however, it should be also noted at this time Phelesmon can not digivolve into Beelzemon.

(<http://en.wikipedia.org/wiki/Phelesmon>, 3rd June 2007)

b) Lexical mistakes

Lexical mistakes include frequent repetitions of vocabulary items (54), which indicate a degree of ineptitude of expression, e.g.:

Air rage

It is considered a problem considering that unlike ground based vehicles, the vehicle is entering altitudes where changes in air pressure can help trigger temporary psychological changes.

(http://en.wikipedia.org/wiki/Air_rage, 1st March 2007)

Oxyd

There are clues on many landscapes: some are helpful, but others are not so helpful.

(<http://en.wikipedia.org/wiki/Oxyd>, 2nd January 2007)

Pharnaces II of Pontus

Mithridates was keen to wage war with the Romans once more, but his son was less keen, and began to plot to remove his father from power.

(http://en.wikipedia.org/wiki/Pharnaces_II_of_Pontus, 21st May 2007)

c) Graphic mistakes

Graphic mistakes comprise orthographic errors, including spelling (17) and punctuation errors (23), e.g.:

Ubon

It is known as Ubon for short. The name mean s "Royal Land Lotus Blossom.

(<http://en.wikipedia.org/wiki/Ubon>, 23rd January 2007)

Unified Energy system

The company itself is over 50% owned by the ruussian state.

(http://en.wikipedia.org/wiki/RAO_UES, 25th March 2007)

Uttarakhand Kranti Dal

The party chose to support the BJP, which was much closer to having a majority compared to the Coongress Party.

(http://en.wikipedia.org/wiki/Uttarakhand_Kranti_Dal, 9th March 2007)

Ube City

Previously a coal mining town the city has developed an effective policy to improve its environment.

(http://en.wikipedia.org/wiki/Ube_City, 21st August 2007)

The rather low presence of graphic mistakes can be attributed to the wiki-mechanism itself and to the use of the so-called “bots”, which automatically correct spelling and punctuation errors.

As the extracts presented above illustrate, the mistakes authors make range from minor errors resulting from negligent editing to mistakes which significantly impair the clarity of the articles.

The instances of deviant definitions prove that the authors collaborating in the creation of articles do not always possess an awareness of the genre. This unawareness can, at times, verge on discourse ineptitude, where the authors instead of contributing to the extension of the informative content of the articles, contribute to a higher level of content unintelligibility.

To conclude, anti-encyclopedic articles represent texts which are functionally inappropriate, which violate not only the criterion of intelligibility and language correctness but, more importantly, the need to adjust the form and content of the articles to a specific communicative situation. The generic and discourse failures of these articles can be attributed to either the author’s unfamiliarity with or a purposeful ignorance of the norms and conventional frameworks associated with the discourse of an encyclopedic article. The authors of the articles quoted above do not reveal an adequate degree of awareness of the genre and of its audience. Therefore, doubts arise whether everyone can become an author of an encyclopedic entry. The consideration of the distinctiveness of the genre seems especially relevant as far as encyclopedic texts are concerned due to their universal character.

6.2. Wikidiscourse transformation areas

The discourse observed in Wikipedia, and the overall shape of Wiki-articles constitute a clear consequence of the process of adjustment of a text to a new context. The discourse of the free encyclopedia clearly reflects the digital, dynamic and global environment in which it is produced, and exhibits the “polyphony of the voices” of its disparate collection of authors. The collaborative character of Wikipedia constitutes a strong determinant in the discourse of this encyclopedia, as the personal styles and idiosyncratic language features of particular authors become apparent.

It should be mentioned that Wikipedia, analogically to most of the publishing houses, provides a style manual for contributors. The manual contains guidelines for article preparation, concerning the stylistic layer of the encyclopedia, capitalization, punctuation, spelling, etc. The aim of the Wikipedia Manual is: “to make the encyclopedia easier to read”, claiming that “one way of presenting information is often just as good as another, but consistency promotes professionalism, simplicity and greater cohesion”, thus advising a consistent application of style and formatting throughout

an article.⁷ Yet, as the analysis proves, the lack of a formal requirement for each contributor to become familiar with those principles and the lack of any formal measures taken in the case of the author's noncompliance with the Manual lead to inconsistency between the articles in nearly every aspect of its discourse.

Moreover, this inconsistency may be ascribed to the lack of formal criteria governing the eligibility to become a contributor. As the founder of Wikipedia states "I don't care if they're a high school kid or a Harvard professor, it's the work that matters."⁸ Though praiseworthy and with great potential, the editorial freedom of Wikipedia, and its openness towards volunteers, somewhat undermines the basic properties of the genre.

Consequently, in comparison with traditional standard print encyclopedias, Wikipedia manifests a departure from the uniform discourse features towards a greater degree of individualization:

Uniformization

→

Individualization

different structure of articles
(based on individual approaches
to layout and content presentation),

different stylistic layers of articles
according to individual discourse
competence level.

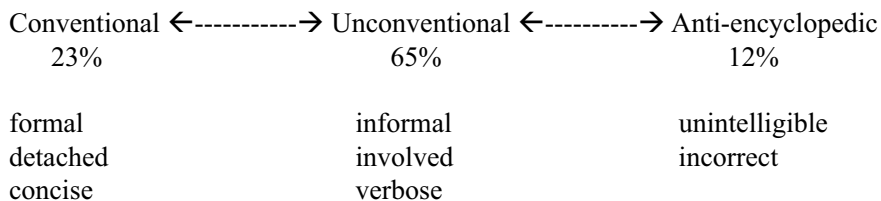
Individualization signals a certain "fragmentation of discursive norms and conventions" to use Fairclough's (1992: 221) terms, which, in the case of Wikipedia, is a clear consequence of the democratisation of the genre. Individualization can be observed in different approaches towards the manner and style of definitions as well as in the incorporation into the articles of divergent content, forms and conventions associated with other text-types. Due to these interferences and modifications, Wiki-articles deviate from the conventional composition of encyclopedic articles, departing from the compositional fixity and clichés associated with the genre.

It is the greater degree of individualization that constitutes the reason for the emergence of the analysed continuum of Wiki-articles. The three major groups of articles arranged along the dimension exhibit the following properties:

⁷ http://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style

⁸ CSpan interview with Jimmy Wales: <http://www.q-and-a.org/Transcript/?ProgramID=1042>

The Continuum of Wiki-entries



The departure from the convention, in the case of unconventional and anti-encyclopedic articles, is manifested on all strata of discourse and can be observed in the modification on textual, syntactic and lexical levels.

The following Table presents the most significant discourse areas affected by the transformation as well as the direction of the change:

STANDARD ENCYCLOPEDIA DISCOURSE	UNCONVENTIONAL WIKIPEDIA DISCOURSE
Homogeneity →	Heterogeneity – presence of foreign genre conventions – citations, narrative elements – internal stylistic syncretism
Compactness →	Verbosity – redundancy – lexical prolixity
Clarity →	Ambiguity – lack of preciseness – restricted cohesion
Formality →	Informality/orality – presence of colloquial lexical and syntactic constructions – use of idiomatic expressions – use of hedges – looser structure of messages
Impersonality →	Personalisation – subjectivity – digressions – involvement, reader-oriented discourse
Competence →	Mixed degree of competence – various levels of awareness of genre conventions – variable degrees of discourse competence on textual, syntactic and lexical levels

Table 1. Wiki-discourse transformation areas

The major defects which disrupt the communicative effect and lead to un-intelligibility of chosen Wikipedia articles concern the following areas:

- discourse level: inadequacy of the texts to the communicative situation, texts stylistically deviant;
- textual level: encompassing incompleteness of description, circularity of definitions, lack of cohesion and coherence, unclear structuring of entries, unclear expression of thoughts;
- syntactic means: deviant word order;
- lexical level: use of contextually improper vocabulary;
- orthographic level: spelling and punctuation mistakes.

The number and diversity of unconventional discourse properties prove that the scope of modifications is widespread. Using Wikipedia as an example illustrates that genre modification may be influenced by an individual realization of the convention, which is in turn influenced by the degree of language competence. As Bhatia (2004) emphasizes, the maintenance of generic conventions may be attained only when all the participants share the knowledge of the genre, that is the rules defining its construction and use. In the case of the free encyclopedia, though, the degree of discourse competence and genre awareness among Wikipedians is uneven, which leads to marked differences in the linguistic layer of the articles. Though traditional paper encyclopedias occasionally contained “little masterpieces of style and content” (Collison 1964: 227), such idiosyncrasy within articles is especially strong in Wikipedia, which is characterized by flexibility, openness and language freedom. The articles involve high-quality compositions as well as those which exhibit the limited language proficiency and genre awareness of their authors. Consequently, Wikipedia offers a collection of texts ranging from an expository presentation of information, through an involved and interactive expression of personal views, to unintelligible and obscure texts.

The departure from a typical encyclopedic formalism, terminological exactitude and conciseness makes Wikipedia different from the universal standard encyclopedias available in a paper format. It needs to be noted, though, that such language freedom makes Wikipedia more accessible to a larger circle of internet users. The increase in the level of informality, the use of colloquial means of expression, a higher interactivity of discourse and a widespread exemplification enhance the readability of the articles, thus facilitating the comprehension of the concepts discussed in the encyclopedia, although naturally, with the exclusion of the incomprehensible entries.

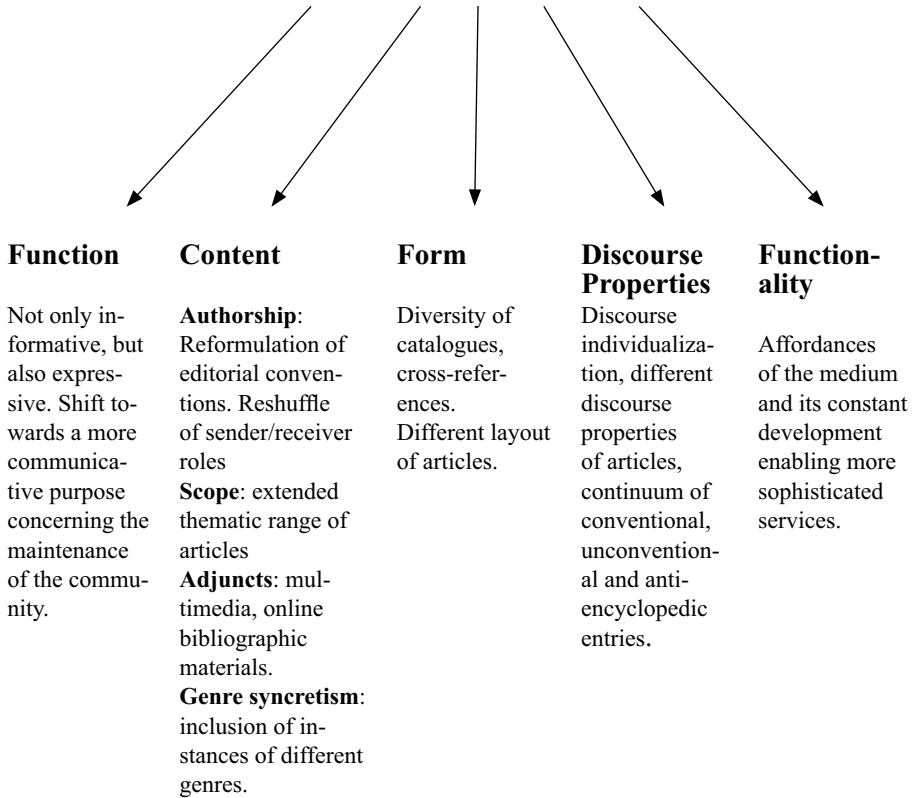
Conclusions

The analysis undertaken allows us to identify the main tendencies in the development of the genre of online encyclopedias. The collection of the analysed web sites proves that encyclopedias have evolved from a mere reproduction of the genre towards highly complex, dynamic and audio-visual resources which fully exploit the potential of the internet. All these innovations illustrate an expansion of the traditional notion of an encyclopedia. Moreover, the interplay of diverse processes affecting the genre has led to the emergence of various sub-types of online encyclopedias characterized by distinctive properties of the generic structure, substantially different from standard print encyclopedic works.

The influence of the medium on the evolution of the genre is best exemplified by the emergence of Wikipedia. The comparison of Wikipedia and print encyclopedias allows us to distinguish a number of similar and a few differentiating features between the two manifestations and to illustrate how the genre of encyclopedia, influenced by social and technological determinants, is being reshaped in a new environment, which constitutes a confirmation of Yates and Orlikowski's (1992) theory concerning the evolution of genres. Changes indicating a departure from the conventions are reflected in the alteration and expansion of the functional, pragmatic, stylistic and structural properties of the genre. The transformations confirm our initial assumptions concerning the extension of conventional components of the genre in the case of Wikipedia.

The following illustration presents the major areas of genre change and evolution:

Genre Transformation Areas



As the areas of transformation illustrate, the contemporary view of an encyclopedia departs from those associated with the beginnings of the genre. The electronic environment in which Wikipedia is created is affected by the considerable dynamism and span of the confrontation between tradition and change. Such a confrontation is clearly observable in the changes affecting the function and content of the free encyclopedia, but also the discourse of Wikipedia, which exhibits features conventionally associated with encyclopedic textual patterns, combined with novel forms, previously absent from the genre.

Therefore, there arises the question regarding the genological affinity of the project. Since Wikipedia and its print antecedents have a number of characteristics in common, as the structure of the free encyclopedia still comprises the genre components of a traditional paper encyclopedia, the free encyclopedia cannot be treated as a novel, spontaneous genre. At the same time, though, judging by the changes introduced to the traditional structure of the genre, it also does not represent a mere reproduction of the genre in a new environment. Its in-between composite structure places Wikipedia among the emergent genres, in which the conventional genre characteristics undergo far reaching changes due to the affordances of the internet technologies.

The merger of the convention with the potential of the medium has both advantageous and unfavourable consequences, which may be summarized as follows:

Advantages	Disadvantages
<ul style="list-style-type: none"> • User-friendliness, availability • Easiness of article composition • Lack of space constraints • New types of information • Multimedia • Elimination of the restrictions of linearity • Availability of external materials • Constant revision • Multi-purpose navigation facilities • Simplified discourse: restricted meta-language, content presented in different styles. 	<ul style="list-style-type: none"> • Lack of guaranteed accuracy and validity • Misinformation • Low number of experts collaborating in the project • Malleability and fluidity of contents • Susceptibility to vandalism • Possibility of copyright violation • Uneven quality of articles • Unintelligibility of some entries.

Table 1. Advantages and disadvantages of Wikipedia as an encyclopedia

Wikipedia clearly exemplifies the phenomenon of genre evolution. It shows that genres develop and change to meet the needs of changing technology and new environments, but also that genres change when individuals take liberties and experiment with the existing conventions.

We might suppose that all the quoted instances exemplifying a change in genre properties are exceptions illustrating a lack of authorial discourse competence. On the other hand, though, it appears that these modifications indicate the course of genre development. Similar tendencies in evolution and analogous changes affecting the genre can already be observed in other encyclopedic sites created on the web. The majority of the encyclopedias are highly interactive, a growing number is created by users and volunteers, and many exhibit tendencies towards genre hybridization, all of which contribute to an expansion of the traditional concept of an encyclopedia. This evolution of the genre is further reflected in the recent decision by the Microsoft to discontinue the publication of the Encarta products. As the reason the editors give the changing nature of reference sources, and also the transformations in the roles, needs and habits of the public: “the category of traditional encyclopedias and reference material has changed. People today seek and consume information in considerably different ways than in years past” (http://encarta.msn.com/guide_page_FAQ/FAQ.html). Similarly, the Britannica Online has recently considered the introduction of the option of a partial contribution, inviting users to send comments, data and ready articles to the editorial board. All these steps taken by the publishers point to their understanding of the transformations occurring in the market, and also their awareness of the need to follow these transformations by implementing new technological solutions.

It is worth noting also the role Wikipedia plays in the democratization of information and science. As Willinsky (2007) observes, Wikipedia constitutes a confirmation of Yochai Benkler’s prophesy that networked information systems will contribute

to the openness of the public sphere. Wikipedia, just like other works of reference, represents “the myth of universal knowledge”, as noted by Jean Noel Jeanneney, the president of Bibliothèque Nationale de France (quoted in Adams 2007: 9). As Adams (*ibid.*: 9) claims, users yearn for an “Infotopia” (a term coined by Cass Sunstein), referring to absolute freedom and democracy of knowledge. Wikipedia clearly illustrates Levy’s (2002) observations that nowadays the acquisition of knowledge falls outside the institutions so far occupied with teaching, live sources of knowledge become widely available and individuals have the opportunity to join virtual communities in which they are given a chance to learn by cooperation and an exchange of information, which, as Adams observes, undermines the whole system regulating science. This supports the claims of Benkler (2002: 15) that “ubiquitous computer communications networks have brought about a dramatic change in the scope, scale, and efficacy of peer production.” The patterns of computer-mediated cooperation exemplified by Wikipedians suggest, thus, the emergence of a new genre, not only of an information resource, but of a collaborative activity (Bryant *et al.* 2005: 9). Moreover, as Holloway *et al.* (2005: 5) observe, through its policy of open-access, the free encyclopedia may offer a possibility to debate over controversial and unsettled matters.

This particular democratization of information and publishing embodied by Wikipedia makes it a unique resource, very different from the reference works we have so far consulted. It makes Wikipedia also one of the most revolutionary projects of our times, representing a new realization of the genre. Wikipedia functions as an encyclopedia in the social consciousness, although there are firm grounds, raised also in this book, for which it should not be treated as an instance of this genre (e.g. inaccuracy, vandalism, expert-rejection). Yet, its social recognition and widespread usage induce us to believe that the project fulfils an important role in the common quest for information. The popularity of Wikipedia, but also the intensive development of similar collaborative projects, require a change in our perception of the function and specificity of reference sources as such. It has become vital to understand “the epistemology of mass collaboration” (Fallis 2009) and to raise users’ awareness concerning the background of such projects and help them develop an ability of a proper exploitation of the information obtained from these sources. What becomes significant at this moment, as various scholars observe (e.g. Fallis *ibid.*, Sanger 2009), is the development of new techniques for retrieving information and verifying the validity of the data.

Bibliography

Sources

Online encyclopedias

A.D.A.M. Medical Encyclopedia: <http://www.nlm.nih.gov/medlineplus/encyclopedia>
Archimedes Laboratory: <http://www.archimedes-lab.org>
ARKive – Images of Life on Earth: <http://www.arkive.org>
Artist's Toolkit Encyclopedia: <http://www.artsconnected.org/toolkit/encyclopedia.html>
Australian War Memorial's Online Encyclopedia: <http://www.awm.gov.au>
Citizendium: http://en.citizendium.org/wiki/Main_Page
Concise Encyclopedia of Economics: <http://www.econlib.org/LIBRARY/CEE.html>
Conservapedia: http://www.conservapedia.com/Main_Page
Cut the Knot: <http://www.cut-the-knot.org/index.shtml>
Digital Universe: <http://www.digitaluniverse.net>
EH.Net Encyclopedia of Economic and Business History: <http://eh.net/encyclopedia>
Electrochemistry Encyclopedia: <http://electrochem.cwru.edu/ed/encycl>
Encyberpedia: <http://www.encyberpedia.com/ency.htm>
Encyclopedia Al-Islam: <http://www.al-islam.org/encyclopedia>
Encyclopedia Astronautica: <http://www.astronautix.com>
Encyclopedia Britannica: <http://www.britannica.com>
Encyclopedia Metallum: <http://www.metal-archives.com/>
Encyclopedia Mythica: <http://www.pantheon.org>
Encyclopedia of Astrobiology: <http://www.daviddarling.info/encyclopedia.html>
Encyclopedia of Astronomy and Astrophysics: <http://eaa.crcpress.com>
Encyclopedia of the Atmospheric Environment: <http://www.ace.mmu.ac.uk/eae/html>
Encyclopedia of Catholicism: <http://www.newadvent.org/cathen/h.htm>
Encyclopedia of Chicago History: <http://www.encyclopedia.chicagohistory.org>
Encyclopedia of Design: <http://www.thedesignencyclopedia.org>
Encyclopedia of Earth: <http://www.eoearth.org>
Encyclopedia of Educational Technology: <http://coe.sdsu.edu/eet>
Encyclopedia of Health: http://www.ajc.com/health/content/shared/adam_index.html
Encyclopedia of Interaction Design: <http://www.interaction-design.org/encyclopedia>
Encyclopedia of Laser Physics and Technology: <http://www.rp-photonics.com>
Encyclopedia of Life Support Systems: <http://greenplanet.eolss.net.aspx>
Encyclopedia of Mathematics: <http://www.mathacademy.com/pr/prime>

Encyclopedia of Mathematics: <http://eom.springer.de>
 Encyclopedia of Mediaeval Ages: <http://www.the-orb.net/encyclo.html>
 Encyclopedia of Philosophy: <http://www.iep.utm.edu>
 Encyclopedia of Plants: <http://www.botany.com/6.sh.html>
 Encyclopedia of Psychology: <http://www.psychology.org/links>
 Encyclopedia of Revolutions: <http://www.ohiou.edu/~chastain>
 Encyclopedia of Star Wars: http://starwars.wikia.com/wiki/Main_Page
 Encyclopedia Smithsonian: http://www.si.edu/Encyclopedia_SI/default.htm
 Encyclopedia of the Sustainable Development: <http://www.ace.mmu.ac.uk/esd.html>
 Encyclopedia of World War I: <http://www.firstworldwar.com/atoz/index.htm>
 Encyc.org Encyclopedia: <http://encyc.org>
 Everything2: <http://www.everything2.com>
 Extrasolar Encyclopedia: <http://exoplanet.eu/biblio.php>
 Fact Monster: <http://www.factmonster.com>
 Fish Base: <http://www.fishbase.org/search.php>
 Free Dictionary: <http://www.thefreedictionary.com>
 Free internet Encyclopedia: <http://www.cam-info.net/enc/micro.a.html>
 Freesoft: <http://www.freesoft.org/CIE>
 GLBTQ Encyclopedia: <http://www.glbtq.com>
 Heart and Stroke Encyclopedia: <http://www.americanheart.org/presenter.jhtml>
 High-Beam Encyclopedia: <http://www.encyclopedia.com>
 Hoax Encyclopedia: <http://antivirus.about.com/od/emailhoaxes/l/blenhoax.htm>
 H2G2: <http://www.bbc.co.uk/dna/h2g2>
 Infoplease: <http://www.infoplease.com/encyclopedia>
 Jewish encyclopedia: <http://www.jewishencyclopedia.com>
 Literary Encyclopedia: <http://www.litencyc.com>
 Love To Know Classic Encyclopedia: <http://www.1911encyclopedia.org>
 Maths is Fun: <http://www.mathsisfun.com>
 Mathworld: <http://mathworld.wolfram.com>
 MSN Encarta: http://encarta.msn.com/artcenter/_browse.html
 Networking Encyclopedia: <http://www.networkworld.com/links/Encyclopedia>
 New Georgia Encyclopedia: <http://www.georgiaencyclopedia.org/nge/Home.jsp>
 New Zealand Peoples Encyclopedia: <http://www.teara.govt.nz/NewZealanders>
 Online Encyclopedia of Integer Sequences: <http://www.research.att.com/~nja/sequences>
 Open Site: <http://open-site.org>
 Panda Security Encyclopedia: http://www.pandasoftware.com/virus_info/encyclopedia
 Photonics Encyclopedia: <http://www.rp-photonics.com/encyclopedia.html>
 Physics Encyclopedia: <http://members.tripod.com/~IgorIvanov/physics>
 Planetmath: <http://planetmath.org/encyclopedia>
 Probert Encyclopedia: <http://www.probert-encyclopaedia.co.uk>
 Reference Encyclopedia: <http://www.reference.com>
 Routledge Encyclopedia of Philosophy: <http://www.rep.routledge.com>
 Scholarpedia: <http://www.scholarpedia.org>
 Small Business Encyclopedia: <http://www.entrepreneur.com/encyclopedia/index.html>
 Stanford Encyclopedia of Philosophy: <http://plato.stanford.edu>
 Storm Encyclopedia: <http://www.weather.com/encyclopedia/winter/types.html>

Techno-web Encyclopedia: <http://www.techweb.com/encyclopedia>
 Toonopedia: <http://www.toonopedia.com>
 Transportation Demand Management Encyclopedia: <http://www.vtpi.org/tadm>
 Uncyclopedia: http://uncyclopedia.org/wiki/Main_Page
 WhatIs Encyclopedia: <http://whatis.techtarget.com>
 Wikipedia – the free encyclopedia: <http://en.wikipedia.org>
 World Book Encyclopedia: <http://www.worldbook.com>

References

- Adams, M. 2007. "The Critical Dictionary and the WikiWorld." *English Today* 90, 23/2, 9–18.
- Alvar, E. 1978. *The Syntactic Structure of English Texts: a Computer-Based Study of Four Kinds of Text in the Brown University Corpus*. Gothenburg: Acta Universitatis Gotheburgensis.
- Arystoteles. 1988. *Retoryka. Poetyka*. Warszawa: PWN.
- Askehave, I., A.E. Nielsen. 2005. "What are the Characteristics of Digital Genres? – Genre Theory from a Multi-modal Perspective." *Proceedings of the 38th Hawaii International Conference on System Sciences*. CDRom Collection. Hawaii: IEEE Press.
- Askehave, I., J.M. Swales. 2001. "Genre Identification and Communicative Purpose: A Problem and a Possible Solution." *Applied Linguistics*, 22/2, 195–212.
- Bachtin, M. 1986. *Estetyka twórczości słownej*. Warszawa: Państwowy Instytut Wydawniczy.
- Bakhtin, M. 1986. *Speech Genres and Other Essays*. Holoquist, M., C. Emerson (ed.). Austin: University of Texas Press.
- Baron, N.S. 2000. *Alphabet to Email: How Written English Evolved and Where it's Heading*. London: Routledge.
- Bardini, T. 1997. "Bridging the Gulfs: from Hypertext to Cyberspace." *Journal of Computer-Mediated Communication*, 3/2. Available online at: <http://jcmc.indiana.edu/vol3/issue2/bardini.html>
- Bartmiński, J. (ed.). 1993. *Encyklopedia kultury polskiej XX w.* Wrocław: Wiedza o Kulturze.
- Bartmiński, J. 1998. "Tekst jako przedmiot tekstologii lingwistycznej." In: Bartmiński, B. Boniecka (ed.). *Tekst. Problemy teoretyczne*. Lublin: WUMCS. (9–25)
- Bauer, L. 2002. "The Illusory Distinction between Lexical and Encyclopedic." Available online at: <http://www.victoria.ac.nz/lals/staff/laurie-bauer.aspx>
- Bauman, M.L. 1999. "The Evolution of internet Genres." *Computers and Composition*, 16, 269–282.
- Baym, N.K. 1995. "The emergence of community in computer-mediated communication." In: S. Jones (ed.). *Cybersociety: Computer-mediated community and communication*. Thousand Oaks: Sage. (138–163)
- Bazerman, C. 1988. *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison: University of Wisconsin Press.
- Beaugrande, R. de, W.U. Dressler. 1990. *Introduction to Text Linguistics*. London and New York: Longman.
- Benkler, Y. 2002. "Coase's Penguin; or, Linux and the Nature of the Firm." *Yale Law Journal*, 112. Available online at: <http://www.benkler.org/CoasesPenguin.html>
- Bergquist, M., J. Ljungberg. 1999. "Genres in Action: Negotiating Genres in Practice." *Proceedings of the 32nd Hawaii International Conference on System Science*. CDRom Collection. Hawaii: IEEE Press.
- Berkenkotter, C., T.N. Huckin. 1995. *Genre Knowledge in Disciplinary Communication*. Hillsdale: Lawrence Erlbaum and Associates.
- Bex, T. 1996. *Variety in Written English*. London: Routledge.

- Bhatia, V.K. 1993. *Analyzing Genre: Language Use in Professional Settings*. London: Longman.
- Bhatia, V.K. 1997. "The Power and Politics of Genre." *World Englishes*, 16/3, 359–371.
- Bhatia, V.K. 1997. "Genre Analysis Today." *Revue Belge de philologie et d'histoire*, 75, 629–652.
- Bhatia, V.K. 2002. "Applied Genre Analysis: Analytical Advances and Pedagogical Procedures." In: Johns, A.M. (ed.). (279–283)
- Bhatia, V.K. 2004. *Worlds of Written Discourse*. London: Continuum.
- Biber, D. 1988. *Variation Across Speech and Writing*. Cambridge: CUP.
- Biber, D. 1989. "A Typology of English Texts." *Linguistics*, 27, 3–43.
- Biber, D. 1995. *Dimensions of Register Variation. A Cross-linguistic Comparison*. Cambridge: CUP.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Boese, E., A. Howe. 2005. "Effects of Web Document Evolution on Genre Classification." *CIKM*, 632–640.
- Bolter, J.D. 2001. *Writing Space: Computers, Hypertext, and the Remediation of Print*. Mahwah: Lawrence Erlbaum, Associates.
- Bradshaw, L.E. 1981. "John Harris's *Lexicon technicum*." In: Kafker, F.A. (ed.). (107–137)
- Breure, L. 2001. *Development of the Genre Concept*. Available online at: <http://www.cs.uu.nl/people/leen/GenreDev/GenreDevelopment.htm>
- Brown J.S., P. Duguid. 1991. "Organizational Learning and Communities of Practice: Toward a Unified View of Working, Learning, and Innovation." *Organization Science*, 2, 40–57.
- Brown, G., G. Yule. 1983. *Discourse Analysis*. Cambridge: CUP.
- Bryant, S., A. Forte, A. Bruckman. 2005. "Becoming a Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia." *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, 1–10. Available online at: <http://portal.acm.org/citation.cfm?id=1099205>
- Buckland, M.K. 2007. "The Digital Difference in Reference Collections." *Journal of Library Administration*, 46/2, 87–100.
- Bucy, E.P. 2004. "Interactivity in Society: Locating an Elusive Concept." *The Information Society*, 20, 373–383.
- Bugajski, M. 1993. *Językoznawstwo normatywne*. Warszawa: WN PWN.
- Burke, S.M. 2003. "The design of Online Lexicons." In: van Sterkenburg, P. (ed.). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins. (240–249)
- Burns, A., C. Coffin. (eds.). 2001. *Analysing English in a Global Context: a Reader*. London: Routledge.
- Burrell, P. 1981. "Pierre Bayle's Dictionnaire historique et critique." In: Kafker, F.A. (ed.). (83–103)
- Bush, V. 1945. "As We May Think." In: Nyce J. M., P. Kahn (eds.). *From Memex to Hypertext* (85–110). Reprinted from the *Atlantic Monthly*, 176/1 (1945), 641–649.
- Capocci, A., P. Servedio, F. Colaiori, L. Buriol, D. Donato, S. Leonardi, G. Calderelli. 2006. "Preferential Attachment in the Growth of Social Networks: the Case of Wikipedia." Available online at: <http://arXiv:physics/0602026v2>
- Carels, P., D. Flory. 1981. "Johann Heinrich Zedler's Universal Lexicon." In: Kafker, F.A. (ed.). (165–169)
- Carr, M. 1997. "Internet Dictionaries and Lexicography." *International Journal of Lexicography*, 10/3, 210–230.
- Carter, R., M. McCarthy. 2006. *Cambridge Grammar of English. A Comprehensive Guide. Spoken and Written English Grammar and Usage*. Cambridge: CUP.
- Castels, M. 2002. *The internet Galaxy. Reflections on the Internet, Business and Society*. Oxford: OUP.

- Chafe, W.L. 1982. "Integration and Involvement in Speaking, Writing, and Oral Literatures." In: Tannen, D. (ed.). *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, N.J: Ablex. (35–53)
- Chafe, W.L. 1985. "Linguistic Differences Produced by Differences between Speaking and Writing." In: Olson D.R., Torrance N., and A. Hildyard (eds.). *Literacy, Language and Learning: The Nature and Consequences of Reading and Writing*. Cambridge: CUP. (105–123)
- Chafe, W.L., J. Danielewicz. 1986. "Properties of Spoken and Written Language." In: Horowitz, R., S.J. Samuels (eds.). *Comprehending Oral and Written Language*. San Diego: Academic Press. (83–113)
- Cherny, L. 1999. *Conversation and Community. Chat in a Virtual World*. Stanford: CSLI Publications.
- Chih-Hua, K. 1999. "The Use of Personal Pronouns: Role Relationships in Scientific Journal Articles." *English for Specific Purposes*, 18/2, 121–138.
- Christie, F., J.R. Martin. 1997. *Genre and Institutions: Social Processes in the Workplace and School*. London: Cassell.
- Collison, R.L. 1964. *Encyclopedias. Their History Throughout the Ages*. London: Hafner.
- Copestake, A. 2006. "Errors in Wikis: New Challenges and New Opportunities – a Discussion Document." Available online at: www.sics.se/jussi/newtext/working_notes/02_copestake.pdf
- Coulson, J., Carr, C.T., Hutchinson, L., Eagle, D., H.M. Peter. (ed.). 1962. *Oxford Illustrated Dictionary*. Oxford: Clarendon Press.
- Coulthard, M., M. Montgomery (eds.). 1981. *Studies in Discourse Analysis*. London: Routledge.
- Coulthard, M. (ed.). 1994. *Advances in Written Text Analysis*. London: Routledge.
- Crowston, K., M. Williams. 1997. "Reproduced and Emergent Genres of Communication on the World Wide Web." *Proceedings of the 30th Annual Hawaii International Conference on System Sciences*. CDROM Collection. Hawaii: IEEE Press.
- Crowston, K., M. Williams. 1999. "The Effects of Linking on Genres of Web Documents." *Proceedings of the 32nd Hawaii International Conference on System Science*. CDROM Collection. Hawaii: IEEE Press. Available online at: <http://crowston.syr.edu/papers/ddgen04.pdf>
- Crowston, K., M. Williams. 2000. "Reproduced and Emergent Genres of Communication on the World Wide Web." *Information Society*, 16/3, 201–216.
- Crowston, K., B. Kwasnik. 2004. "A Framework for a Facetted Classification for Genres: Addressing Issues of Multidimensionality." CDROM Collection. *Proceedings of 37th Hawaii International Conference on Systems Science*. Hawaii: IEEE Press.
- Crystal, D., D. Davy. 1969. *Investigating English Style*. London: Longman.
- Crystal, D. 2001. *Language and the Internet*. Cambridge: CUP.
- Cukier, W., Cody, S., E. Nesselroth. 2006. "Genres of Spam: Expectations and Deceptions." CDROM Collection. *Proceedings of the 39th Hawaii International Conference on System Science*. Hawaii: IEEE Press.
- Cutting, J. 2000. *Analysing the Language of Discourse Communities*. Oxford: Elsevier Science.
- Dalby, A. 2007. "Wikipedias on the Language Map of the World." *English Today* 90, 23/2, 3–8.
- Damaso, J., C. Cotter, 2007. "UrbanDictionary.com." *English Today* 90, 23/2, 19–26.
- Danet, B. 2001. *Cyberplay: Communicating Online*. Oxford: Berg.
- Darnton, R. 1979. *The Business of Enlightenment. A Publishing History of the Encyclopedie 1775–1800*. Cambridge, MA, London: The Belknap Press.
- Davis, B. H., J. P. Brewer. 1997. *Electronic Discourse: Linguistic Individuals in Virtual Space*. NY: State University of New York Press.
- Dąbrowska, A. 1991. "Wartościowanie w wybranych hasłach encyklopedycznych." In: Puzynina J., J. Anusiewicz. (eds.). *Język a kultura. Wartości w języku i tekście*, vol. 3. Wrocław: WUWr. (85–100)

- Desilets, A., S. Paquet, N. Vinson. 2005. "Are Wikis Usable?" *Proceedings of the 2005 international symposium on Wikis*. Available online at: www.wikisym.org/ws2005/proceedings/paper-01.pdf
- Devitt, A. 1993. "Generalizing about Genre: New Conceptions of an Old Concept." *College Composition and Communication*, 44/4, 573–586.
- Dewdney, N., VanEss-Dykema, C., R. MacMillan. 2001. "Classification of Genres in Text." Available online at: <http://www.elsnet.org/km2001/dewdnew.pdf>
- Dewe, J., Bretan, I., J. Karlgren. 1998. "Assembling a Balanced Corpus from the internet." *Proceedings of the 11th Nordic Computational Linguistics Conference*, 100–107.
- Dillon, A., B. Gushrowski. 1996. "Genres and the Web: Is the Personal Home Page the First Uniquely Digital Genre?" *Journal of the American Society for Information Science*, 51/2, 202–205.
- Dobrzyńska, T. (ed.). 1992. *Typy tekstów*. Warszawa: IBL PAN.
- Duszak, A. 1997. "Analyzing Digressiveness in Polish Academic Texts." In: Duszak, A. (ed.). *Cultures and Styles of Academic Discourse*. Berlin, New York: Mouton de Gruyter. (323–341)
- Duszak, A. 1998. *Tekst, dyskurs, komunikacja międzykulturowa*. Warszawa: Wydawnictwo Naukowe PWN.
- Eco, U. 2002. "Nowe środki masowego przekazu a przyszłość książki." In: Hopfinger, M. (ed.). *Nowe media w komunikacji społecznej XX wieku: antologia*. Warszawa: Oficyna Naukowa. (536–554)
- Eco, U. 2003. *O literaturze*. Warszawa: Muza.
- The Encyclopedia of Language and Linguistics*. Vol. 3. 1994. Asher, R. (ed.). Oxford: Pergamon Press.
- Erickson, T. 1997. "Social Interaction on the Net: Virtual Community as Participatory Genre." CDROM Collection. *Proceedings of the 30th Annual Hawaii International Conference on System Sciences*. Hawaii: IEEE Press.
- Erickson, T. 1999. "Rhyme and Punishment: The Creation and Enforcement of Conventions in an On-Line Participatory Limerick Genre." *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences*. CDROM Collection. Hawaii: IEEE Press.
- Erickson, T. 2000. "Making Sense of Computer-Mediated Communication: Conversations as Genres, CMC Systems as Genre Ecologies." *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. CDROM Collection. Hawaii: IEEE Press. Available online at: http://www.pliant.org/personal/Tom_Erickson/genreEcologies.html
- Fairclough, N. 1992. *Discourse and Social Change*. Cambridge: Polity Press.
- Fairclough, N. 1995. *Critical Discourse Analysis: The Critical Study of Language*. London and New York: Longman.
- Fallis, D. 2009. "Introduction: The Epistemology of Mass Collaboration." *Episteme*, 6/1, 1–7.
- Farrell, T. 1977. "Literacy, the Basics and All That Jazz." *College English*, 538, 445–459
- Ferrara, K., H. Brunner, and G. Whittemore. 1991. "Interactive Written Discourse as an Emergent Register." *Written Communication* 8, 8–34.
- Floetttum, K. Torodd, K., T. Dahl. 2007. "We Now Report on...Versus Let us Now See How: Author Roles and Interaction with Readers in Research Articles." In: Hyland, K., M. Bondi (eds.). (203–224)
- Freedman, A., P. Medway (eds.). 1994. *Genre and the New Rhetoric*. London: Taylor and Francis.
- Frost, W. (ed.). 2005. *The Reference Collection – from the Shelf to the Web*. New York: The Haworth Press.
- Fuchs, U. 2006. "Freie Inhalte? - Idee und Realisierung am Beispiel der Wikipedia." In: Ohly P., Siegler-schmidt J., C. Swertz (eds.). *Wissensorganisation und Verantwortung: Gesellschaftliche, ökonomische und technische Aspekte. Proceedings der 9. Tagung der Deutschen Sektion der Internationalen Gesellschaft für Wissensorganisation, Duisburg 5.–7. November 2004*. Würzburg: Ergon Verlag. (192–198)

- Furuta, R, C.C. Marshall. 1995. "Genre as Reflection of Technology in the World Wide Web." *Proceedings of the International Workshop on Hypermedia Design*. Berlin: Springer. (203–214)
- Gajda, S. 1982. *Podstawy badań stylistycznych nad językiem naukowym*. Warszawa: Wydawnictwo Naukowe.
- Gajda, S. 1990. *Współczesna polszczyzna naukowa. Język czy żargon?* Opole: Instytut Śląski.
- Gajda, S. 1993. "Gatunkowe wzorce wypowiedzi." In: Bartmiński, J. (ed.). (283–304)
- Gajda, S. (ed.). 1999. *Dyskurs naukowy – tradycja i zmiana*. Opole: WUO.
- Gann, K. 2007. "Sand Castles of Knowledge." *ArtsJournal.com*. Available online at: http://www.artsjournal.com/postclassic/2007/05/sand_castles_of_knowledge.html
- Gasson, S. 2004. *Qualitative IS Research Methods*. Available online at: <http://www.cis.drexel.edu/faculty/gasson/Research.html#b2>
- Georgakopoulou, A., D. Goutsos. 1999. *Discourse Analysis. An Introduction*. Edinburgh: EUP.
- Giddens, A. 1984. *The Constitution of Society: Outline of the Theory of Structure*. Berkeley: University of California Press.
- Ghadessy, M. (ed.). 1988. *Registers of Written English: Situational Factors and Linguistic Features*. London: Pinter.
- Ghadessy, M. (ed.). 1993. *Register Analysis: Theory and Practice*. London: Pinter.
- Goban-Klas, T. 2004. *Media i komunikowanie masowe: Teorie i analizy radia, prasy, telewizji i internetu*. Warszawa: Wydawnictwo Naukowe PWN.
- Görlach, M. 2004. *Text Types and the History of English*. Berlin, New York: Mouton de Gruyter.
- Goldstein, J., R.E. Sabin. 2006. "Using speech acts to categorize email and identify email genres." *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*. CDROM Collection. Hawaii: IEEE Press.
- Gotti, M., D.S. Giannoni. (eds.). 2006. *New Trends in Specialized Discourse Analysis*. Frankfurt a. M.: Peter Lang.
- Greenbaum, S. 1969. *Studies in English Adverbial Usage*. London: Longman.
- Greer, J., D. Mensing. 2004. "The Evolution of Online Newspapers: A longitudinal content analysis, 1997–2003." *Newspaper Research Journal*, 25/2, 98–112.
- Grice, P. 1978. "Further Notes on Logic and Conversation." In Cole, P. (ed.). *Syntax and Semantics: Pragmatics*, vol. 9. New York: Academic Press. (183–97)
- Grogan, D. 1987. *Grogan's Case Studies in Reference Work*. Vol. 2 *Encyclopedias, Yearbooks, Directories and Statistical Sources*. London: Clive Bingley.
- Gruszczyński, W. 2001. "Czaty w sieci." *Dialog*, 2, 137–145.
- Grybosiowa, A. 2003. "Cudzysłów jako wyznacznik języka w prasie." In: Grybosiowa, A. (ed.). *Język wtopiony w rzeczywistość*. Katowice: WUŚ. (123–127)
- Grzenia, J. 2006. *Komunikacja językowa w Internecie*. Warszawa: Wydawnictwo Naukowe PWN.
- Gumperz, J. 1982. *Discourse Strategies*. Cambridge: CUP.
- Gygi, K. 1990. "Recognizing the Symptoms of Hypertext...and What to Do about It." In: Laurel, B. (ed.). *Computers as Theatre*. Reading, MA: Addison-Wesley, (279–298).
- Ha, L., E.L. James. 1998. "Interactivity Reexamined: A Baseline Analysis of Early Business Web Site." *Journal of Broadcasting & Electronic Media*, 42/4, 457–474.
- Halavais, A., D. Lackaff. 2008. "An Analysis of Topical Coverage of Wikipedia." *Journal of Computer-Mediated Communication*, 13, 429–440.
- Halliday, M.A.K. 1978. *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- Halliday, M.A.K. 1985. *Spoken and written language*. Oxford: OUP.

- Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*. London: Edward Arnold.
- Halliday, M.A.K., R. Hasan. 1976. *Cohesion in English*. Cambridge: CUP.
- Halliday, M.A.K., R. Hasan. 1989. *Language, Context and Text: Aspects of Language in a Social Semiotic Perspective*. Oxford: OUP.
- Hanks, P. 2001. "The Probable and the Possible: Lexicography in the Age of the internet." *Studies in Lexicography* 11.1, 7–36.
- Hartmann, R. 2005. "Pure or Hybrid? The Development of Mixed Dictionary Genres." *Facta Universitatis*, 3/2, 193–208.
- Herring, S. (ed.). 1996. *Computer Mediated Communication: Linguistic, Social and Cross-cultural Perspectives*. Amsterdam: John Benjamins.
- Herring, S., Scheidt, L.A., Bonus, S., E. Wright. 2004. "Bridging the Gap: A Genre Analysis of Weblogs." *Proceedings of 37th Hawaii International Conference on Systems Science*. CDROM Collection. Hawaii: IEEE Press.
- Hoey, M. 2001. *Textual Interaction: an Introduction to Written Discourse Analysis*. London: Routledge.
- Holloway, T., M. Bozicevic, K. Borner. 2005. "Analyzing and Visualizing the Semantic Coverage of Wikipedia and Its Authors." Special issue on *Understanding Complex Systems*. Available online at: <http://arxiv.org/abs/cs.IR/0512085>
- Horowitz, R., S.J. Samuels (eds.). 1987. *Comprehending Oral and Written Language*. San Diego and London: Academic Press.
- Hunston, S., G. Thompson. (eds.). 2000. *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: OUP.
- Hyland, K. 1998. *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins.
- Hyland, K. 2001. "Bringing in the Reader: Addressee Features in Academic Articles." *Written Communication*, 18/4, 549–574.
- Hyland, K. 2002. "Genre: Language, Context, and Literacy." *Annual Review of Applied Linguistics*, 22, 113–135.
- Hyland, K. 2005. "Stance and Engagement: a Model of Interaction in Academic Discourse." *Discourse Studies*, 7/2, 173–192.
- Hyland, K. 2007. "Applying a Gloss: Exemplifying and Reformulating in Academic Discourse." *Applied Linguistics*, 28/2, 266–285.
- Hyland, K., Bondi, M. (eds.). 2007. *Academic Discourse across Disciplines*. Frankfurt a. M.: Peter Lang.
- Hymes, D. 1974. *Foundations in Sociolinguistics: An Ethnographic Approach*. Philadelphia, PA: University of Pennsylvania Press.
- Jackson, M. 1997. "Assessing the Structure of Communication on the World Wide Web." *Journal of Computer-Mediated Communication*, 3/1. Available online at: <http://www.nicoladoering.de/Hogrefe/jackson.htm>
- Jackson, H. 2002. *Lexicography: an Introduction*. London: Routledge.
- Jacobson, F. 1995. "Encyclopedias." In: Bopp, R.E. (ed.). *Reference and Information Services: an Introduction*. Englewood: Libraries Unlimited. (433–467)
- Jasińska, M. 2001. "Językowy savoir-vivre w internecie." In: Habrajska, G. (ed.). *Język w komunikacji*, vol. III. Łódź: Wydawnictwo Wyższej Szkoły Humanistyczno-Ekonomicznej. (119–125)
- Jean-Pol, M. 2006. "Gemeinsam Wissen konstruieren: am Beispiel der Wikipedia." In: Klebl, M., M. Köck (eds.). *Projekte und Perspektiven im Studium Digitale. Medienpädagogik*, 3. Berlin: LIT. (157–164)
- Jensen, J.F. 1998. "Interactivity. Tracking a New Concept in Media and Communication Studies." *Nordicom Review*, 19, 185–204.

- Johns, A.M. 1997. *Text, Role, and Context. Developing Academic Literacies*. Cambridge: CUP.
- Johns, A.M. (ed.). 2002. *Genre in the Classroom. Multiple perspectives*. London: Lawrence Erlbaum Associates.
- Jones, S. 1998. *Doing Internet Research*. Chicago: Sage.
- Jones, S. 2000. *The internet and Society*. Cambridge: Polity Press.
- Kafker, F.A. (ed.). 1981. *Notable Encyclopedias of the Seventeenth and Eighteenth Centuries: Nine Predecessors of the Encyclopedie*. Banbury, Oxfordshire: Cheney&Sons Ltd.
- Karlgren, J., D. Cutting. 1994. "Recognizing Text Genres with Simple Metrics Using Discriminant Analysis." *Proceedings of the 15th International Conference on Computational Linguistics*, 2, 1071–1075.
- Katz, W. 1982. *Introduction to Reference Work. Vol. 1. Basic Information Sources*. New York, London: McGraw-Hill.
- Katz, W. 1998. *Cuneiform to Computer. A History of Reference Sources*. London: Scarecrow Press.
- Katz, W. 2003. "Reference Books from Cuneiform to Computer." In: Hartmann, R.R.K. (ed.). *Lexicography: critical concepts*, vol. III. London: Routledge, (405–421).
- Keen, A. 2007. *Kult amatora. Jak internet niszczy kulturę*. Warszawa: WAIP.
- Kennedy, A., M. Shepherd. 2005. "Automatic Identification of Home Pages on the Web." *Proceedings of 37th Hawaii International Conference on Systems Science*. CDROM Collection. Hawaii: IEEE Press.
- Kenning, M. 2006. "Evolving Concepts and Moving Targets: Communicative Competence and the Mediation of Communication." *International Journal of Applied Linguistics* 16/3 363–388.
- Kessler, B., Nunberg, G., H. Schutze. 1997. "Automatic Detection of Text Genre." In: Cohen, P.R., W. Wahlster (eds.). *Proceedings of the Thirty-Fifth Annual Meeting of the Association for Computational Linguistics and Eight Conference of the European Chapter of the Association for Computational Linguistics*. Somerset, NJ: Association for Computational Linguistics (32–38).
- Kiousis, S. 2002. "Interactivity: a Concept Explication." *New Media and Society*, 4, 355–383.
- Kister, K. 1981. *Encyclopedia Buying Guide*. London: R.R. Bowker Company.
- Kling, R., E. Callahan. 2003. "Typology of E-journals." *Annual Review of Information Science and Technology*, 37, 127–177.
- Kress, G., T. van Leeuwen. 1996. *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Kress, G. 1998. "Visual and Verbal Modes of Representation in Electronically Mediated Communication." In: Snyder, D. (ed.). *Page to Screen: Taking Literacy into the Electronic Era*. London: Routledge.
- Kuznetsov, S. 2006. "Motivations of Contributors to Wikipedia." *ACM SIGCAS Computers and Society*, 36/2, 1–7.
- Lachowicz, D. 1981. "On the Use of the Passive Voice for Objectivity, Author Responsibility and Hedging in EST." *Science of Science*, 2/2, 105–114.
- Lakoff, G. 1972. "Hedges: A Study of Meaning Criteria and the Logic of Fuzzy Concepts." In: Peranteau, P., Levi, J., G. Phares (eds.). *Papers from the Eighth Regional Meeting of Chicago Linguistic Society*. Chicago: Chicago University Press, (183–228).
- Lan, L. 2005. "The Growing Prosperity of On-line Dictionaries." *English Today* 83, 21/3, 16–21.
- Landau, S. 2001. "What is a Dictionary?" In: Landau, S. (ed.). *Dictionaries: the Art and Craft of Lexicography*. Cambridge: CUP (6–42).
- Landow, G.P. 1992. *Hypertext: The Convergence of Contemporary Critical Theory and Technology*. Baltimore: Johns Hopkins University Press.
- Lanham, R. 1993. *The Electronic Word: Democracy, Technology and the Arts*. Chicago: The University of Chicago Press.
- Library Terms and Definitions <http://www.lib.uoguelph.ca/help/research/terms.htm>

- Levy, P. 2002. "Drugi potop." In: Hopfinger, M. (ed.). (373–390)
- Lyons, J. 1977. *Semantics*. Cambridge: CUP.
- Lyda, A. 2007. *Concessive Relation in Spoken Discourse*. Katowice: WUŚ.
- Mainguy, G. 2006. "Wikipedia and Science Publishing. Has the Time Come to End the Liaisons Dange-reuses?" Paper presented at 3rd NATO-UNESCO Advanced Research Workshop Science Education, Balatonfüred. Available online at: <http://www.nyex.info/dokumentumok/mainguy2.pdf>
- Marszałek, L. 1986. *Edytorstwo publikacji naukowych*. Warszawa: Wydawnictwo Naukowe.
- Martin, J.R. 2001. "Language, Register and Genre." In: Burns, A., C. Coffin. (eds.). (150–166)
- Martin, J.R., D. Rose. 2002. *Working with Discourse: Meaning beyond the Clause*. London: Continuum.
- McArthur, T. 1986. *Worlds of Reference. Lexicography, Learning and Language from the Clay Tablet to the Computer*. Cambridge: CUP.
- McArthur, T. 2003. "What Then is Reference Science?" In: Hartmann, R.R.K. (ed.). *Lexicography: critical concepts*, vol. III. London: Routledge (422–428).
- McCarthy, M. 1991. *Discourse Analysis for Language Teachers*. Cambridge: CUP.
- McHenry, R. 2003. "Whatever Happened to Encyclopedic Style?" *The Chronicle of Higher Education*, 49/25, 13–14.
- McHenry, R. 2004. "The Faith-Based Encyclopedia." *TCS Daily Online*. Available online at: <http://www.tcsdaily.com/article.aspx?id=111504A>
- McLuhan, M. 1964. *Understanding Media: the Extensions of Man*. Cambridge, Mass.: MIT Press.
- Meyer zu Eissen, S., B. Stein. 2004. "Genre Classification of Web Pages. User Study and Feasibility Analysis." In: Biundo, S., Fruehwirth, T., G. Palm (eds.). *KI 2004: Advances in Artificial Intelligence*. Berlin: Springer-Verlag. (256–269)
- Mielczarek 1972. *Z zagadnień leksykografii encyklopedycznej*. Warszawa: Wydawnictwo Naukowe.
- Miller, C. 1984. "Genre as Social Action." In: Freedman, A., P. Medway (ed.). *Genre and the New Rhetoric*. London: Taylor and Francis. (23–42)
- Mizrach, S. 1997. "From Orality to Teleliteracy." Available online at: <http://www.fiu.edu/~mizrachs/orality.htm>
- Moon, R. 1987. "The Analysis of Meaning." In: Sinclair, J. *Looking up*. Harlow, Essex: Longman. (86–103)
- Morse, J.M. 2005. "Reference Publishing in the Age of Also." In: Frost, W. (ed.). (69–83)
- Möller, E. 2006. *Die Heimliche Medienrevolution – Wie Weblogs, Wikis und freie Software die Welt verändern*. Hannover: Heinz Heize.
- Murray, D.E. 2000. "Protean Communication: The Language of Computer-Mediated Communication." *TESOL Quarterly* 34/3, 397–421.
- Nesi, H. 1999. "A User's Guide to Electronic Dictionaries for Language Learners." *International Journal of Lexicography*, 12/1, 56–66.
- The New Encyclopedia Britannica* (EB). vol. 18. 1990. Goetz, P.W. (ed.). Chicago: Encyclopedia Britannica, Inc.
- Niemann, H. 2006. "Was alle angeht, können nur alle lösen – Über die weltweite Anwendung des Kritischen Rationalismus in Wiki-Projekten." *Aufklärung und Kritik*, 1, 67–93.
- Nunan, D. 1993. *Introducing Discourse Analysis*. London: Penguin.
- Nystrand, M. 1982. *What Writers Know. The Language, Process, and Structure of Written Discourse*. New York: Academic Press.
- Oblak, T. 2005. "The Lack of Interactivity and Hypertextuality in Online Media". *The International Journal for Communication Studies*, 67/1, 87–106.

- Online Dictionary of Library and Information Science (ODLIS) Available at: <http://lu.com/odlis/index.cfm>
- Olkiewicz, J. 1988. *Od A do Z, czyli o encyklopediach i encyklopedystach*. Warszawa: Ludowa Spółdzielnia Wydawnicza.
- Ong, W. 2002. *Orality and Literacy*. London: Routledge.
- Orlikowski, W. J., J. Yates. 1994. "Genre Repertoire: The Structuring of Communicative Practices in Organizations." *Administrative Science Quarterly*, 39, 541–574.
- Orlikowski, W. J., J. Yates. 1998. "Genre Systems: Structuring Interaction through Communicative Norms." MIT Sloan School of Management. Available online at: <http://ccs.mit.edu/papers/CCSWP205>
- Orlowski, A. 2005. "Wikipedia Science 31% More Cronky than Britannica's." *The Register*, December 14. Available online at: http://www.theregister.co.uk/2005/12/16/wikipedia_britannica_science_comparison
- Paltridge, B. 1994. "Genre Analysis and the Identification of Textual Boundaries." *Applied Linguistics*, 15, 288–299.
- Paltridge, B. 1995. "Working with Genre: A Pragmatic Perspective." *Journal of Pragmatics*, 24, 393–406.
- Paltridge, B. 2002. "Genre, Text Type, and the English for Academic Purposes (EAP) Classroom." In: Johns, A.M. (ed.). (73–90)
- Pang, A. 1998. "The Work of the Encyclopedia in the Age of Electronic Reproduction." *First Monday*, 3/9. Available at: http://www.firstmonday.org/issues/issue3_9/pang
- Piotrowski, T. 2001. *Zrozumieć leksykografię*. Warszawa: Wydawnictwo Naukowe.
- Pullen, J. 2005. "Freedom of Information." Available online at: <http://magazine.continental.com/072005/content/advance/idea-makers.jsp>
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Rafaeli, S. 1988. "Interactivity: From New Media to Communication." *Sage Annual Review of Communication Research*, 16, 110–134.
- Rafaeli, S., F. Sudweeks. 1997. "Networked Interactivity." *Journal of Computer-Mediated Communication* 2/4. Available online at: www.ascusc.org/jcmc/vol2/issue4/rafaeli.sudweeks.html
- Rehm, G. 2006. "Hypertext Types and Markup Languages." In: Metzging, D., A. Witt (ed.). *Linguistic Modelling of Information and Markup Languages*. Berlin: Springer.
- Rheingold, H. 2000. *The Virtual Community: Homesteading on the Electronic Frontier* (revised edition). Cambridge, Mass.: The MIT Press.
- Roberts, G. 1998. "The Home Page as Genre: Narrative Approach." *Proceedings of the 31st Hawaii International Conference on System Science*. CDROM Collection. Hawaii: IEEE Press.
- Rosenzweig, R. 2006. "Can History Be Open Source? Wikipedia and the Future of the Past." *The Journal of American History*, 93/1, 117–146.
- Roussinov, D., Crowston, K., Nilan, N., Kwasnik, B., Cai, J., X. Liu. 2001. "Genre Based Navigation on the Web." *Proceedings of the 34th Hawaii International Conference on System Science*. CDROM Collection. Hawaii: IEEE Press.
- Russel, T.M. (ed.). 1997a. *The Encyclopaedic Dictionary in the Eighteenth Century: Architecture, Arts and Crafts. Volume Two - Ephraim Chambers Cyclopaedia*. Aldershot: Ashgate Publishing.
- Russel, T.M. (ed.). 1997b. *The Encyclopaedic Dictionary in the Eighteenth Century: Architecture, Arts and Crafts. Volume Five. A Society of Gentlemen Encyclopaedia Britannica*. Aldershot: Ashgate Publishing.

- Ryan, T., Field, R., L. Olfman. 2002. "Homepage genre dimensionality." *Proceedings of the Eighth Americas Conference on Information System*, 1116–1128.
- Sanger, L.M. 2009. "The Fate of Expertise after Wikipedia." *Episteme*, 6/1, 52–73.
- Santini, M. 2006. "Interpreting Genre Evolution on the Web." *EALC 2006 Workshop*. Available online at: http://www.sics.se/jussi/newtext/working_notes/06_santini.pdf
- Santini, M. 2007. "Characterizing Genres of Web Pages: Genre Hybridism and Individualization." *Proceedings of the 40th Hawaii International Conference on System Science*. CDROM Collection. Hawaii: IEEE Press.
- Schiff, S. 2006. "Know it All. Can Wikipedia Conquer Expertise?" *New Yorker*, July 31. Available online at: http://www.newyorker.com/archive/2006/07/31/060731fa_fact
- Schiffrin, D. 1987. *Discourse Markers*. Cambridge: CUP.
- Schmid-Isler, S. 2000. "The Language of Digital Genres – a Semiotic Investigation of Style and Iconology on the World Wide Web." *Proceedings of the 33rd Hawaii International Conference on System Science*. CDROM Collection. Hawaii: IEEE Press.
- Schryver, de G.M. 2003. "Lexicographers' Dreams in the Electronic-Dictionary Age." *International Journal of Lexicography*, 16/2, 144–199.
- Searle, J. 1969. *Speech Acts*. Cambridge: CUP.
- Searle, J.R. 1976. "A Classification of Illocutionary Acts." *Language in Society*, 5/1, 1–24.
- Shepherd, M., C. Watters. 1998. "The Evolution of Cyberggenres." *Proceedings of the 31st Hawaii International Conference on System Science*. CDROM Collection. Hawaii: IEEE Press.
- Shepherd, M., C. Watters. 1999. "The Functionality Attribute of Cyberggenres." *Proceedings of the 32nd Hawaii International Conference on System Science*.
- Shepherd, M., C. Watters, A. Kennedy. 2004. "Cyberggenre: Automatic Identification of Home Pages on the Web." *Journal of Web Engineering*, 3/3–4, 236–251.
- Sieńko, M. 2002. *Człowiek w pajęczynie: internet jako zjawisko kulturowe*. Wrocław: Atut.
- Sinclair, J. 1988. "Compressed English." In: Ghadessy, M. (ed.). (130–136)
- Smith, F. 1985. *Reading without Nonsense*. New York: Teachers College Press.
- Smith, M., P. Kollock. 1999. *Communities in Cyberspace*. London: Routledge.
- Snyder, I. (ed.). 2002. *Silicon Literacies: Communication, Innovation and Education in the Electronic Age*. London: Routledge.
- Soukup, C. 2000. "Building a Theory of Multi-media CMC." *New Media & Society*, 2/4, 407–425.
- Starzec A. 1999. *Współczesna polszczyzna popularnonaukowa*. Opole: WUO.
- Stevens, N. 1991. "Evaluating Reference Books in Theory and Practice." In: Katz, W. (ed.). *Reference and Information Services*. London: The Scarecrow Press. (401–411)
- Stewart, P. 2002. "The Encyclopedie On-line." In: Brewer, D., J. Candler Hayes (eds.). *Using the Encyclopedie. Ways of knowing, Ways of Reading*. Oxford: Voltaire Foundation (173–185).
- Stockwell, F. 2001. *A History of Information Storage and Retrieval*. North Carolina and London: McFarland and Company.
- Stromer-Galley, J. 2004. "Interactivity-as-Product and Interactivity-as-Process." *The Information Society*, 20, 391–394.
- Stvilia, B., M. Twidale, L. Gasser, L. Smith. 2005. "Information Quality Discussions in Wikipedia." In: Hawamdeh, S. (ed.). *Knowledge Management: Nurturing Culture, Innovation and Technology – Proceedings of the 2005 International Conference on Knowledge Management*. Charlotte, NC: World Scientific Publishing Company (101–113). Available online at: <http://mailer.fsu.edu/bstvilia/papers/qualWiki.pdf>

- Sudweeks, F., M. McLaughlin, and S. Rafaeli (eds.). 1998. *Network and Netplay: Virtual Groups on the internet*. Menlo Park, CA: MIT Press.
- Swales, J. 1990. *Genre Analysis*. Cambridge: CUP.
- Swales, J. 1993. "Genre and Engagement." *Revue Belge de philologie et d'histoire*, 71, 687–698.
- Swales, J. 1998. "Consider This: The Role of Imperatives in Scholarly Writing." *Applied Linguistics*, 19/1, 97–121.
- Swales, J. 2004. *Research Genres: Explorations and Applications*. Cambridge: CUP.
- Tannen, D. 1982. *Spoken and Written Language: Exploring Orality and Literacy*. Norwood, NJ: Ablex.
- Tereszkiewicz, A. 2006. "Analiza gatunkowa encyklopedii internetowej Wikipedia." *Biuletyn Polskiego Towarzystwa Językoznawczego*, LXII: 105–115.
- Tereszkiewicz, A. (2008) "Wikipedia as an Encyclopedia – Analysis of a Survey." *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 125: 179–192.
- Todorov, T. 1990. *Genres in Discourse*. Cambridge: CUP.
- Toms, E., D. Campbell. 1999. "Genre ad Interface Metaphor: Exploiting Form and Function in Digital Environments." *Proceedings of the 32nd Hawaii International Conference on System Sciences*. CDROM Collection. Hawaii: IEEE Press.
- Urbańczyk, S. 1991. *Słowniki i encyklopedie. Ich rodzaje i użyteczność*. Kraków–Katowice: Wydawnictwo Katowickie.
- van Dijk, T.A. 1998. *Discourse as Structure and Process*. London: Sage.
- Vaughan, M.W., A. Dillon. 1998. "The role of Genre in Shaping our Understanding of Digital Documents." *Proceedings of 61st Annual Meeting of the American Society for Information Science*. Medford NJ: Information Today Inc. (559–566)
- Ventola, E. 1987. *The Structure of Social Interaction. A Systemic Approach to the Semiotics of Service Encounters*. London: Frances Pinter Publishers.
- Viegas, F. 2007. "The Visual Side of Wikipedia." *Proceedings of the 40th Hawaii International Conference on System Science*. CDROM Collection. Hawaii: IEEE Press.
- Viegas, F. M. Wattenberg, J. Kriss, F. van Ham. 2007. "Talk Before You Type: Coordination in Wikipedia." *Proceedings of the 40th Hawaii International Conference on System Sciences*. CDROM Collection. Hawaii: IEEE Press.
- Vold, E.T. 2006. "Epistemic Modality Markers in Research Articles: a Cross-linguistic and Cross-disciplinary Study." *International Journal of Applied Linguistics*, 16/1, 61–87.
- Wallace, D.P., C. Van Fleet. 2005. "The Democratization of Information?" *Reference and User Services Quarterly*, 45/2, 100–103.
- Warschauer, M. 1999. *Electronic Literacies: Language, Culture and Power in Online Education*. Mahwah, NJ: Lawrence Erlbaum.
- Waters, N., L. 2007. "Why You Can't Cite Wikipedia in My Class." *Communications of the ACM*, 50/9, 15–17.
- Watters, C.R., M.A. Shepherd. 1997. "The Digital Broadsheet: An Evolving Genre." *Proceedings of the 30th Hawaii International Conference on System Science*. CDROM Collection. Hawaii: IEEE Press.
- Wells, H.G. 1938. "World Encyclopedia." Reprinted in: Kochen, M. (ed.). 1967. *The Growth of Knowledge. Readings on Organization and Retrieval of Information*. New York: John Wiley and Sons Inc. (11–22)
- Wielka Encyklopedia Powszechna (WEP)* Vol. 8. 2002. Wojnowski, J. (ed.). Warszawa: Wydawnictwo Naukowe.
- Wierzbicka, A. 1983. "Genry mowy." In: Dobrzyńska, T., E. Janus (ed.). *Tekst i zdanie*. Wrocław: Ossolineum. (125–137)

- Wilkoń, A. 2002. *Spójność i struktura tekstu*. Kraków: Universitas.
- Willinsky, J. 2007. "What Open Access Research Can Do for Wikipedia." *First Monday*, 12/3. Available online at: http://firstmonday.org/issues/issue12_3/willinsky/index.html
- Witosz, B. 2001. "Gatunek-sporny (?) problem współczesnej refleksji tekstologicznej." *Teksty drugie*, 5, 67–85.
- Witosz, B. 2005. *Genologia lingwistyczna. Zarys problematyki*. Katowice: WUŚ.
- Yates, J., W.J. Orlikowski. 1992. "Genres of Organizational Communication: A Structural Approach to Studying Communication and Media." *Academy of Management Review*, 17/2, 299–326.
- Yates, J., Orlikowski, W. J., K. Okamura. 1999. "Explicit and Implicit Structuring of Genres in Electronic Communication: Reinforcement and Change of Social Interaction." *Organizational Science*, 10/1, 83–103.
- Yates, J., Orlikowski, W. J., J. Rennecker. 1997. "Collaborative Genres for Collaboration: Genre Systems in Digital Media." *Proceedings of the 30th Hawaii International Conference on System Sciences*. CDROM Collection. Hawaii. IEEE Press.
- Yates, S.J., T.R. Sumner. 1997. "Digital Genres and the New Burden of Fixity." *Proceedings of the 29th Hawaii International Conference on System Sciences*. CDROM Collection. Hawaii. IEEE Press.
- Yeo, R. 2001. *Encyclopaedic Visions. Scientific Dictionaries and Enlightenment Culture*. Cambridge. CUP.
- Yolton, J.W., Porter, R., Rogers, P, B.M. Stafford. (eds.). 1991. *The Blackwell Companion to the Enlightenment*. Oxford: Blackwell.
- Yoshioka, T., Herman, G., Yates, J., W. Orlikowski. 2001. "Genre taxonomy: A Knowledge Repository of Communicative Actions." *ACM Transactions on Information Systems*, 19/4, 431–456.
- Zaśko-Zielińska, M. 2002. *Przez okno świadomości. Gatunki mowy w świadomości użytkowników języka*. Wrocław: WUWr.
- Zlatić, V., M. Boxicević, H. Stefancic, M. Domazet. 2006. "Wikipedias: Collaborative Web-based Encyclopedias as Complex Networks." Retrieved from: <http://arXiv:physics/0802149v3> on 17th March 2007.
- Żydek-Bednarczuk, U. 2004. "Tekst w Internecie i jego wyznaczniki." In: Kita, M., J. Grzenia (ed.). *Dialog a nowe media*. Katowice: WUŚ. (11–19)
- http://commons.wikimedia.org/wiki/Commons:Community_Portal
- <http://commons.wikimedia.org/wiki/Category:Animations>
- <http://en.wikipedia.org>
- http://en.wikipedia.org/wiki/Category:Spoken_articles
- http://en.wikipedia.org/wiki/John_Smith_%28disambiguation%29
- http://en.wikipedia.org/wiki/Wikipedia:Citing_sources
- <http://en.wikipedia.org/wiki/Wikipedia:Copyrights>
- http://en.wikipedia.org/wiki/Wikipedia:List_of_policies
- http://en.wikipedia.org/wiki/Wikipedia:Multilingual_statistics
- <http://en.wikipedia.org/wiki/Wikipedia:Statistics>
- http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Fact_and_Reference_Check
- http://new-arch.rp.pl/artikul/722116_Wikipedia_troche_mniej_wolna.html
- http://personalweb.about.com/od/wikihostingandsoftware/Wiki_Sites_Wikia_Wikicities_etc.htm
- http://wikimediafoundation.org/wiki/Collaborative_Video
- http://wikimediafoundation.org/wiki/Wikipedia_Invites_Users_to_Take_Part_in_Open%2C_Collaborative_Video_Experiment
- <http://www.gnu.org>

http://www.kaltura.com/devwiki/index.php/Main_Page

<http://www.lib.uoguelph.ca/help/research/terms.htm>

<http://www.q-and-a.org/Transcript/?ProgramID=1042>

http://www.theregister.co.uk/2005/12/16/wikipedia_britannica_science_comparison

http://www.wikia.com/wiki/List_of_Wikia

REDAKTOR
Renata Włodek

SKŁAD I ŁAMANIE
Katarzyna Mróz

KOREKTA
Dorota Bednarska

Wydawnictwo Uniwersytetu Jagiellońskiego
Redakcja: ul. Michałowskiego 9/2, 31-126 Kraków
tel. 12-631-18-81, 12-631-18-82, fax 12-631-18-83